Correspondence

A Subspace Identification Extension to the Phase Correlation Method

William Scott Hoge

***Abstract*—The phase correlation method (PCM) 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 noninteger pixel displacement identification without interpolation, robustness to noise, and limited computational complexity make this approach a very attractive extension of the PCM. In addition, this approach is shown to be complementary with other subpixel phase correlation based techniques.**

***Index Terms*—Phase correlation method, subpixel image registration, SVD.**

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 interimage 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 [2], [3]. Thus, image registration using phase correlation can be applied directly to the raw MRI data before the spatial images are reconstructed.

The PCM 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 (2-D) functions, A and B, that are related by a simple translational shift, the elements of the Fourier transform of B, denoted $\mathcal{B}$, are related to $\mathcal{A}$ by

$$\mathcal{B}(k,l) = \mathcal{A}(k,l)\exp\{-j(ka+lb)\} \tag{1}$$

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.

Manuscript received April 20, 2002; revised September 18, 2002. This work was supported by the National Institutes of Health (NIH) under Grant (T32–PA-00-103). The Associate Editor responsible for coordinating the review of this paper and recommending its publication was A. Manduca.

The author is with the Harvard Medical School and Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115 USA (e-mail: shoge@ieee.org).

Digital Object Identifier 10.1109/TMI.2002.808359

To identify a and b using the PCM approach, one computes a normalized cross power spectrum between $\mathcal{A}$ and $\mathcal{B}$ to identify the *phase correlation matrix*

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

Once computed, the approach cited in the literature [5], [6] is to compute the inverse Fourier transform of $\mathcal{Q}$. If the two functions under comparison were in fact continuous, then this representation would be a delta function,

$$Q(x,y) = \delta(x-a, y-b) \tag{3}$$

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

However, in the case of images (which are 2-D functions sampled on a discrete grid), Q will only display a delta-like function if a and b are integers. Noninteger 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 Section II-B.

To identify noninteger 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 (LSF) to a 2-D 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 LSF 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 $\mathcal{Q}$ is in fact a rank-one matrix. Each element in $\mathcal{Q}$ can be separated as $\mathcal{Q}(k,l) = \exp\{-jka\}\exp\{-jlb\}$. This allows the definition of two vectors

$$q_a(k) = \exp\{-jka\} \text{ and } q_b(l) = \exp\{+jlb\} \tag{4}$$

and the phase correlation matrix can be rewritten as $\mathcal{Q} = q_a q_b^H$. This allows one to rewrite (1) in matrix notation as

$$\mathcal{B} = \left(q_a q_b^H\right) \circ \mathcal{A} \tag{5}$$

where $\{\cdot\}^H$ denotes a complex-conjugate transpose, and $\circ$ indicates an element-by-element product, also known as the Schur or Hadamard product [9].

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, $\mathcal{Q}$. A straightforward approach to finding the dominant rank-one subspace of $\mathcal{Q}$ 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 noninteger translational motion over a large range.

To identify the linear phase coefficients in each of the right and left dominant singular vectors of $\mathcal{Q}$, a LSF to the unwrapped phase component of the dominant singular vectors is 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 $R\ [\ \mu\ c\ ]^T = \mathrm{unwrap}\{\angle v\}$ 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, μ and c are the slope and abscissa of the fitted line, respectively. This system is then solved to give

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

The slope of the fitted line, μ, 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 $\mathcal{Q}$ 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 2-D phase unwrapping of the matrix $\mathcal{Q}$. 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 PCM (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 nonideal 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] and [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, $\mathcal{Q}$, to restrict the spectrum components corrupted by aliasing from the shift estimation. This mask captures the components of the data matrix $\mathcal{A}$ with magnitude larger than a given threshold α 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 dc 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 2-D 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 Medical Systems (Fairfield, CT) Signa L $\times$ 1.5-Tesla MRI scanner. The 256×256 pixel image covers a 16 cm^2 FOV. Thus, the extent of each pixel is 0.0625 mm square. The echo repetition time for each of the acquired images was $\mathrm{TR} = 500$ ms, 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 Section II-A; and the method given in [7], labeled here as "2-D LSF". This distinction is given because both

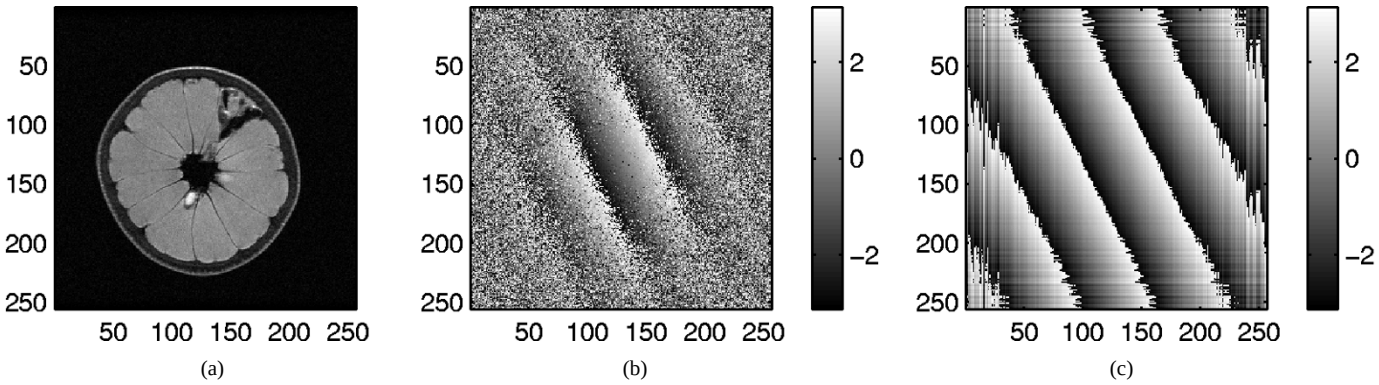

Fig. 1. Acquired MR image example. (a) Grapefruit MR image acquired using an FSE sequence (Image 1 of Table I). (b) Phase correlation matrix, $\mathcal{Q}$, for grapefruit images 1 and 2. (c) Rank-one approximation constructed from the dominant singular vectors of $\mathcal{Q}$. The dc component of the frequency domain coordinate system is at the center of the phase correlation matrices.

TABLE I
POSITION OF EACH ACQUIRED GRAPEFRUIT IMAGE RELATIVE TO THE SERIES LANDMARK

Image:	1	2	3	4	5
Vertical Pos. (in mm):	3.0	1.5	0.0	-1.5	-1.5
Horizontal Pos. (in mm):	0.0	2.5	5.0	2.7	7.5

of the translation estimate methods use frequency masking. The primary difference is that SIE-PCM estimates the shifts separately across a large noninteger range, whereas 2-D LSF first uses a coarse integer registration (based here on the SIE-PCM estimate), then uses a 2-D LSF for subpixel 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, $\mathcal{Q}$.

Fig. 1 shows various stages of the SIE-PCM method for one image pair's phase correlation matrix, $\mathcal{Q}$. Inspection of Fig. 1(b) shows significant noise in $\mathcal{Q}$ directly attributable to the low SNR in the compared images. Fig. 1(c) shows the rank-one approximation of $\mathcal{Q}$ 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 dc) 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. And 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 2-D LDF method of [7]. The measures shown are [12] the actual estimate error, $\mathrm{AE}(A_1, A_2) = \|A_1 - A_2\|_F$, and the relative estimate error, $\mathrm{RE}(A_1, A_2) = \|A_1 - A_2\|_F / \|A_1\|_F$, with the Frobenius norm defined as $\|A\|_F = \left[\sum_{ij} a_{ij}^2\right]^{-1/2}$.

Fig. 2 displays the relative error (RE) for each possible comparison of the images listed in Table I. Note that for the exception of two cases,

TABLE II
REGISTRATION RESULTS BETWEEN GRAPEFRUIT IMAGES 1 AND 2. THE QUALITY OF THE REGISTRATION IS MEASURED USING THE AE AND RE MEASURES

	Vertical	Horizontal
physical translation	-1.5 mm	-2.5 mm
image distance (in pixels)	-2.4	-4.0
SIE-PCM estimate (in pixels)	-2.0613	-4.0241
SIE-PCM estimate error	-0.3387	0.0241
2D LSF estimate (in pixels)	-2.0627	-4.0252
2D LSF estimate error	-0.3373	0.0252

registration error	AE	RE
expected shift:	2518.0261	0.2060
2D LSF estimated shift:	2378.9403	0.1946
SIE-PCM estimated shift:	2378.8575	0.1946

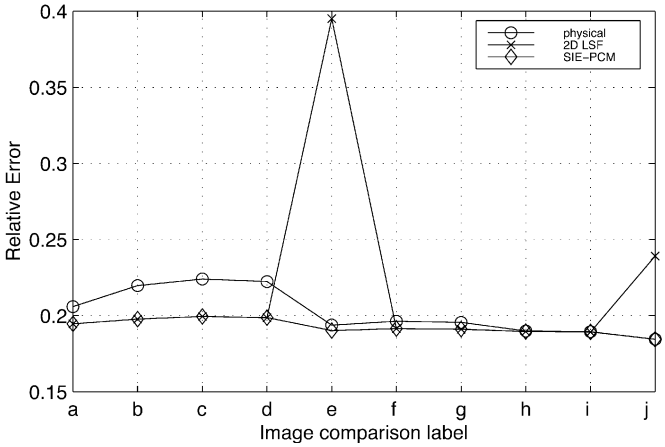

Fig. 2. Comparison of relative error across registration methods. The registration methods shown are (1) knowledge of the "physical" shift, (2) the 2-D LSF method of Stone, *et al.*, [7], and (3) the SIE-PCM method. Each data point in the figure shows the relative error after registering one grapefruit image to a second grapefruit image. The image pairs registered are (a) 1:2, (b) 1:3, (c) 1:4, (d) 1:5 (e) 2:3, (f) 2:4, (g) 2:5, (h) 3:4, (i) 3:5, and (j) 4:5.

both shift estimation methods provide more accurate registration than knowledge of the prescribed shift. The two anomalous 2-D LSF results are attributed to the fact that the 2-D LSF method must unwrap the phase data in two dimensions before estimating the fitted plane parameters. In the anomalous cases shown, there was insufficient unwrapping of the noisy phase correlation matrix before translation estimation. Because phase unwrapping need only occur along one dimension in the SIE-PCM approach, this example clearly demonstrates the utility of separately estimating the shift parameters.

IV. Summary

This paper presents a method to identify bulk translational motion between two images. The method is an extension of the popular phase correlation technique, with the added advantage of noninteger translational shift identification over a wide range of potential shift values. The primary feature is that an SVD factorization of the phase correlation matrix is used to separate the horizontal and vertical components of the translation. A linear fit to the phase of these separate components is then performed in order to identify the magnitude of the translation.

The advantage of the method is threefold. First, noninteger or subpixel displacement can be directly determined without spatial domain interpolation or two-step course-then-fine registration. Second, by separating the horizontal and vertical displacement estimation, phase unwrapping of the data is reduced to one dimension. Finally, this approach is complementary to other subpixel image registration approaches. Section II-B presented frequency masking to compensate for aliasing and boundary effects, complementing the method presented in [7]. One could also potentially combine this approach with the method described in [4].

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.

Acknowledgment

The author would like to thank Dr. E. L. Miller, Dr. D. H. Brooks, and Dr. H. Foroosh for their comments in regards to this paper. He would also like to thank the anonymous reviewers for their instructive and informative comments that strengthened this paper.

References

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

[2] G. A. Wright, "Magnetic resonance imaging," *IEEE Signal Processing Mag.*, vol. 14, pp. 56–66, Jan. 1997.

[3] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan, *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. New York: Wiley, 1999.

[4] H. Foroosh, J. B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Trans. Image Processing*, vol. 11, pp. 188–200, Mar. 2002.

[5] A. M. Tekalp, *Digital Video Processing*. Upper Saddle River, NJ: Prentice-Hall, 1995.

[6] C. Stiller and J. Konrad, "Estimating motion in image sequences," *IEEE Signal Processing Mag.*, vol. 16, pp. 70–91, July 1999.

[7] H. S. Stone, M. Orchard, E.-C. Chang, and S. Martucci, "A fast direct Fourier-based algorithm for subpixel registration of images," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 2235–2243, Oct. 2001.

[8] H. Stone, M. Orchard, and E.-C. Chang, "Subpixel registration of images," in *Rec. 33rd Asilomar Conf. Signals, Systems, and Computers*, vol. 2, 1999, pp. 1446–1452.

[9] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. New York: Cambridge Univ. Press, 1991.

[10] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, pp. 51–83, Jan. 1978.

[11] S. Alliney and C. Morandi, "Digital image registration using projections," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, no. PAMI-2, pp. 222–233, Mar. 1986.

[12] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*. New York: Academic, 1990.

[13] B. S. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE Trans. Image Processing*, vol. 5, pp. 1266–1271, Aug. 1996.