ESTIMATING SUB-PIXEL SHIFTS DIRECTLY FROM THE PHASE DIFFERENCE

Murat Balci *Hassan Foroosh*

School of Computer Science, University of Central Florida, Orlando, FL, 32816-3262
{balci,foroosh}@cs.ucf.edu

ABSTRACT

In this paper, we establish the exact relationship between the continuous and the discrete phase-difference of two shifted images, and show that their discrete phase difference is a 2-dimensional sawtooth signal. As a result, the exact shifts between two images can be determined to sub-pixel accuracy by counting the number of cycles of the phase difference matrix along each frequency axis. The sub-pixel portion is represented by a fraction of a cycle corresponding to the non-integer part of the shift. The problem is formulated as an over-determined system of equations, and is solved by imposing a regularity constraint, using the method of Generalized Cross Validation (GCV).

1. MOTIVATION

In this paper, we are motivated by applications that require registration at sub-pixel accuracy such as examination of same patient Magnetic Resonance Imaging (MRI) data in a clinical setting [2], or super-resolution from multiple frames [6, 10, 8]. We are particularly motivated by sensor modalities that provide directly the Fourier spectrum of the field of view, e.g. MRI, and Synthetic Aperture Radar (SAR).

Among existing sub-pixel registration techniques Fourier domain methods [3, 5, 11, 4] are an important class of registration techniques that have gained popularity due to their remarkable accuracy. These methods are in general variations of the original phase correlation method [7]. For instance, in [3] Foroosh et al. demonstrated how the method can be extended to sub-pixel accuracy, in [11] stone et al. investigated the effect of aliasing error, and in [5] Hoge describes how the shift parameters can be decoupled in the dominant left and right eigenvectors of the phase correlation matrix.

This paper revisits these concepts and shows that the analytic model for the discrete phase difference matrix is a 2-dimensional sawtooth signal. This in particular leads to a simple solution directly from the phase difference matrix in the form of counting the number of cycles in the sawtooth signal, using an overdetermined system of equations. The overdetermined nature of the formulation allows for handling random noise.

This work was partially supported by ONR grant #N00014-04-1-0512, and Sun Microsystem's grant #EDUD-7824-030482-US.

2. CONTINUOUS MODEL

Consider two continuous signals $f_1(x,y)$ and $f_2(x,y) = f_1(x-x_o, y-y_o)$, whose Fourier transforms exist (i.e. are square-integrable). Their cross power spectrum (also known as the phase correlation) is then given by

$$\hat{c}(u,v) = \frac{\hat{f}_1 \hat{f}_2^*}{|\hat{f}_1 \hat{f}_2^*|} = \exp(-ix_o u - iy_o v) \qquad (1)$$

where the hat sign indicates the Fourier transform, and the asterisk stands for the complex conjugate.

In other words, spatial translations lead to linear phase differences between the two signals along each frequency axis, i.e. [1]

$$\hat{\varphi}(u,v) = \angle \hat{c}(u,v) = x_o u + y_o v \qquad (2)$$

which is a planar surface through the origin. As a result, the spatial translations can be determined by inverse transforming $\hat{c}(u,v)$, which leads a Dirac delta function centered at (x_o, y_o).

A discrete interpretation of this continuous-domain result is used in practice for image registration [1, 7], which yields very good results, and was also extended to sub-pixel registration in [3]. However, for some sensors such as MRI and SAR that provide directly the Fourier spectrum of the field of view, it would be interesting from both practical and computational points of view to estimate the shifts directly in the Fourier domain. A practical solution for this problem was first proposed by Hoge [5]. His method requires a subspace approximation by Singular Value Decomposition (SVD), and unwrapping of the dominant left and right eigenvectors. We will show below that both unwrapping and subspace approximation may be unncessary steps, since we derive an analytic model of the phase difference matrix and solve the problem by fitting the model to the noisy data.

[1] Throughout this paper, we call $\exp(-ix_o u - iy_o v)$ the phase correlation function and $\hat{\varphi}(u,v) = \angle \hat{c}(u,v)$ the phase differnce function. Similarly, their discrete counterparts are referred to as the discrete phase correlation matrix, and the discrete phase difference matrix, respectively.

0-7803-9134-9/05/$20.00 ©2005 IEEE

3. DISCRETE MODEL

While the continuous model provides insight to the problem, it can be misleading for the discrete data. To demonstrate this, let us examine the phase difference function in (2). Inverse transforming this phase angle would yield

$$\varphi(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_0u + y_ov)\text{exp}(iux + ivy)dudv \quad (3)$$

$$= -ix_o\frac{d\delta(x)}{dx} - iy_o\frac{d\delta(y)}{dy} \quad (4)$$

where the derivatives are understood in the distributional sense [13].

From bandlimited sampling theory and (4), it follows that the spatial domain representation of a component of the discrete phase matrix is given by

$$\varphi_{kl} = -i\frac{x_o}{\pi k}\left(2\text{sinc}\frac{\pi k}{x_o} - 2\cos\frac{\pi k}{x_o}\right) - i\frac{y_o}{\pi l}\left(2\text{sinc}\frac{\pi l}{y_o} - 2\cos\frac{\pi l}{y_o}\right) \quad (5)$$

On the other hand, note that

$$i\varphi_{kl} = \frac{x_o^2}{\pi k^2}\left(2\frac{\pi k}{x_o}\cos\frac{\pi k}{x_o} - 2\sin\frac{\pi k}{x_o}\right)$$

$$+\pi\frac{y_o^2}{\pi l^2}\left(2\frac{\pi l}{y_o}2\cos\frac{\pi l}{y_o} - 2\sin\frac{\pi l}{y_o}\right) \quad (6)$$

$$= x_o\frac{2}{2\pi/x_o}\int_{-\frac{\pi}{x_o}}^{\frac{\pi}{x_o}} u\sin ku\ \text{d}u + y_o\frac{2}{2\pi/y_o}\int_{-\frac{\pi}{y_o}}^{\frac{\pi}{y_o}} v\sin lv\ \text{d}v \quad (7)$$

It can also be verified that

$$x_o\frac{2}{2\pi/x_o}\int_{-\frac{\pi}{x_o}}^{\frac{\pi}{x_o}} u\cos ku\text{d}u + y_o\frac{2}{2\pi/y_o}\int_{-\frac{\pi}{y_o}}^{\frac{\pi}{y_o}} v\cos lv\text{d}v = 0 \quad (8)$$

and similarly

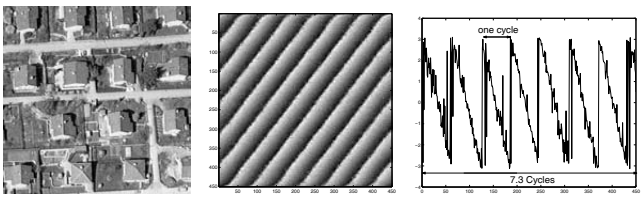
$$x_o\frac{2}{2\pi/x_o}\int_{-\frac{\pi}{x_o}}^{\frac{\pi}{x_o}} u\text{d}u + y_o\frac{2}{2\pi/y_o}\int_{-\frac{\pi}{y_o}}^{\frac{\pi}{y_o}} v\text{d}v = 0 \quad (9)$$

From (7), (8), and (9), and using the definition of the Discrete Fourier Transform (DFT) based on Fourier series [9], it follows immediately upon substituting $u = n\frac{2\pi}{N}$ and $v = m\frac{2\pi}{M}$ that φ_{kl} is a DFT coefficient of the following discrete periodic signal

$$\hat{\varphi}_{mn} = \frac{2\pi}{N}\left(x_on + j\frac{N}{x_o}\right) + \frac{2\pi}{M}\left(my_o + k\frac{M}{y_o}\right) \quad (10)$$

where j and k are arbitrary integers.

In other words the discrete phase difference matrix for a pair of shifted images is given by $\mathbf{\Phi} = [\hat{\varphi}_{mn}]$, where $m = 0, ..., M-1$, and $n = 0, ..., N-1$. This is a discrete 2D periodic sawtooth signal as opposed to the continuous phase difference function in (2), which is a plane through the origin. Figure 1 shows examples of noisy discrete phase difference matrices. It can be verified from (10) that the periods of this discrete sawtooth signal along the two axis are given by $\left(\frac{2\pi}{x_o}, \frac{2\pi}{y_o}\right)$. In other words, there are x_o repeated cycles along each row of the phase difference matrix, where x_o may or may not be an integer. When x_o is not an integer, the number of repeated cycles in a row is given by the integer part of x_o plus a fraction of a cycle defined by the non-integer portion of x_o. A similar argument applies to the columns of $\mathbf{\Phi}$. This process of counting the number of cycles along the rows and columns of the phase matrix is essentially all that is required to determine the shifts. In the next two sections, we will design a linear estimator for this problem.



(a) (b) (c)

Fig. 1. (a) One of the two images being registered, (b) noisy sawtooth phase matrix corresponding to shifts of (7.3,5.6) pixels, (c) one row of the phase difference matrix.

4. PROPOSED METHOD

As indicated above, the key to solve the problem is to find how many cycles of the discrete sawtooth phase difference fit in the range $[0, 2\pi]$ along each frequency axis. The number of cycles i.e. x_o and y_o may or may not be integer values, and are given by

$$x_o = \frac{\text{cycles}}{2\pi} = \frac{N}{2\pi}\frac{d\mathbf{\Phi}(m,n)}{dn} \quad (11)$$

$$y_o = \frac{\text{cycles}}{2\pi} = \frac{M}{2\pi}\frac{d\mathbf{\Phi}(m,n)}{dm} \quad (12)$$

These are by definition the instantaneous frequencies. Due to noise and other sources of error, however, counting the number of cycles per 2π using equations (11) and (12) may lead to inaccurate results. To overcome this problem, we exploit the fact that a total of $M \times N$ data points are available for regression. Therefore, an accurate solution can be obtained by fitting the parameters of the analytic model to the huge data points.

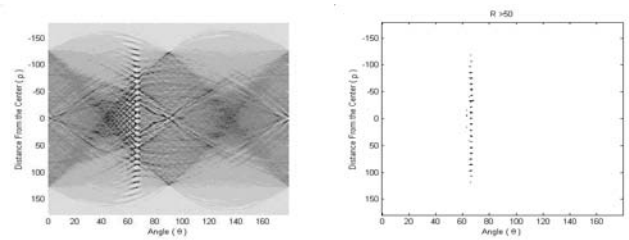
The counting process visually involves detecting the set of points where the sawtooth signal takes a given value (e.g. zero) along each row or column. This can be implemented in practice by seting $j = 0$ and finding all points for which the phase difference matrix has a zero crossing. Algebraically this is equivalent to solving

$$n\cos\theta + m\sin\theta + \rho = 0 \quad (13)$$

where $\tan\theta = \frac{Ny_o}{Mx_o}$ and $\rho = k\frac{M}{y_o}\sin\theta$.

This shows that the locus of the set of points for which the phase difference is zero is a family of lines given by

$$\frac{\rho}{\cos\theta}x_o + \frac{\rho}{\sin\theta}y_o = k(M+N) \quad (14)$$



(a) (b)

Fig. 2. (a) The Hough transform of the phase discontinuities, (b) Peaks detected by thresholding.

Each line itself is parameterized by the angle θ and its distance ρ from the center frequency (i.e. the origin of the frequency domain). This is an ideal setup for Hough transform, which maps these lines to a parameter space of (θ, ρ). As can be verified above, θ remains invariant among all lines and ρ varies as integer multiples of some other invariant parameter, i.e. $\rho = k\rho_o$. Therefore in the Hough-transform domain, we expect to see a set of peak values situated at equal distances from each other, and parallel to the ρ- axis Figure 2 shows an example of the Hough transform of the discontinuities of the phase difference matrix of two shifted images, where the peaks can be clearly identified.

Each peak point in the Hough-transform domain provides one linear constraint on x_o and y_o. Given a total of t such peak values, we can construct a system of linear equations of the form

$$\mathbf{Ar} = \mathbf{b} \tag{15}$$

$$\text{where} \quad \mathbf{A} = \begin{bmatrix} \frac{\rho_1}{\cos\theta_1} & \frac{\rho_1}{\sin\theta_1} \\ \vdots & \vdots \\ \frac{\rho_t}{\cos\theta_t} & \frac{\rho_t}{\sin\theta_t} \end{bmatrix} \tag{16}$$

$\mathbf{r} = [x_o \;\; y_o]^T$, and $\mathbf{b} = (M+N)[k_1 \cdots k_t]^T$.

The norm-regularized solution [12] to this over-determined system of equations is given by

$$\mathbf{r}_{\text{opt}} = (\mathbf{A}^T\mathbf{A} + \lambda\mathbf{L}^T\mathbf{L})^{-1}\mathbf{A}^T\mathbf{b} \tag{17}$$

where λ is a regularization parameter, and $\mathbf{L}$ is such that

$$\mathbf{L}^T\mathbf{L} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \tag{18}$$

This choice of $\mathbf{L}$ [12] implies that our *a priori* belief is that our solution should be identical over all equations in the system, i.e. the equations should be consistent with each other. The optimal choice for the regularization parameter is given by the method of GCV, which amounts to minimizing

$$\text{GCV}(\lambda) = \frac{\|(\mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{L}^T\mathbf{L})^{-1}\mathbf{A}^T)\mathbf{b}\|^2}{(\text{tr}(\mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{L}^T\mathbf{L})^{-1}\mathbf{A}^T))^2} \tag{19}$$

with respect to λ, where $\text{tr}(\cdot)$ is the trace of a matrix. A first order approximation to the optimal value of λ is given by [2]:

$$\lambda^* = \frac{\sigma_1 \sum_{j=2}^{t} s_j^2}{(t-1)s_1^2 - \sum_{j=2}^{t} s_j^2} \tag{20}$$

where σ_1 is the dominant eignevalue of $\mathbf{AL}^{-1}\mathbf{L}^{-T}\mathbf{A}^T$, and s_j are the components of the vector $\mathbf{V}^T\mathbf{b} = [s_1 \; s_2 \; \ldots \; s_t]^T$, where $\mathbf{V}$ is the matrix of the eigenvectors of $\mathbf{AL}^{-1}\mathbf{L}^{-T}\mathbf{A}^T$.

Using (17) and (20), we can compute the optimal solutions for x_o and y_o.

5. RESULTS AND DISCUSSION

We applied the technique to both synthetic and real data. For the synthetic set, we generated some shifted images with additive zero-mean Gaussian noise. The input image and a typical phase difference matrix are shown in Fig. 3. We then compared the variation of the average SNR in our technique with that of Hoge over 50 independent trials, while varying the standard deviation of the noise in the range $[0, 5]$. Results are shown in Figure 4. Overall the two methods have similar performance, although it seems that on average, over the range of the noise levels of this experimentation, Hoge's method was performing slightly better at the lower half of the noise interval, and our method was performing slightly better at the upper half of the interval.

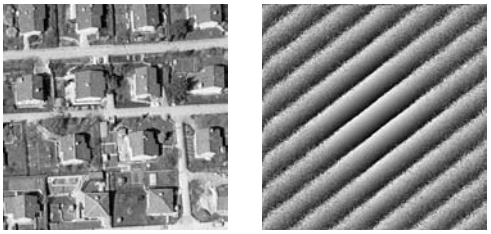


Fig. 3. Simulation with additive noise: left: image used to synthesize shifted noisy pairs, right: a typical phase difference matrix for a noisy pair.

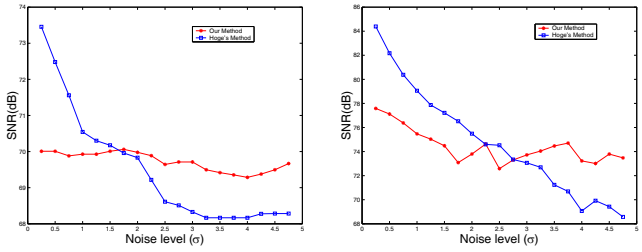
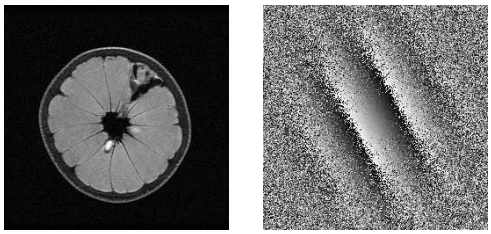


Fig. 4. Average SNR over 50 independent trials for the standard deviation of noise varying in $[0, 5]$: left shift along x-axis, right shift along y-axis.

We also applied our technique to a set of real MRI data, used by Hoge in his experimentation [5]. Figure 5 shows

[2]The proof is omitted due to lack of space.

an MRI image from this data set, and the phase difference matrix for one pair. The phase difference matrix is noisy and highly affected by aliasing. We repeated Hoge's exact experimental procedure. Table 1 summarizes the shift parameters obtained by our approach compared to those obtained by physical calibration, Stone et al. and Hoge. As described in [5], due to manual displacement errors the physical calibration may be more erroneous than the computed methods. Therefore similar to [5], we compared the different algorithms using the relative error. Results are shown in Fig. 6. Overall, given the very small variations of the relative error, the methods seem to be comparable in terms of performance.



(a) (b)

Fig. 5. (a) An image from the MRI data set (Courtesy of W.S. Hoge), (b) discrete phase difference matrix of the first pair in the data set.

Image Pairs	Physical (in pixels)	Stone et al.	Hoge	Proposed Method
(1,2)	(-2.40, -4.00)	(-2.06,-5.97)	(-2.05,-4.02)	(-2.11,-4.00)
(1,3)	(-4.80,-8.00)	(-4.28,-9.99)	(-4.26,-8.01)	(-3.90,-7.49)
(1,4)	(-7.20,-4.32)	(-6.61,-5.67)	(-6.61,-4.33)	(-6.22,-3.93)
(1,5)	(-7.20,-12.00)	(-6.62,-13.97)	(-6.61,-12.03)	(-6.39,-11.42)
(2,3)	(-2.40,-4.00)	(-5.86,-3.85)	(-2.21,-3.98)	(-2.18,-3.87)
(2,4)	(-4.80,-0.32)	(-4.55, -1.71)	(-4.54,-0.32)	(-4.16,-0.31)
(2,5)	(-4.80,-8.00)	(-4.56,-10.00)	(-4.55,-8.01)	(-4.13,-7.73)
(3,4)	(-2.40,3.68)	(-2.34,2.34)	(-2.33,3.66)	(-2.34,3.55)
(3,5)	(-2.40,-4.00)	(-2.34,-5.97)	(-2.34,-4.03)	(-2.49,-3.83)
(4,5)	(0.00,-7.68)	(-0.81,-8.05)	(-0.01,-7.69)	(-0.03,-7.84)

Table 1. Pairwise registration of the MRI data set.

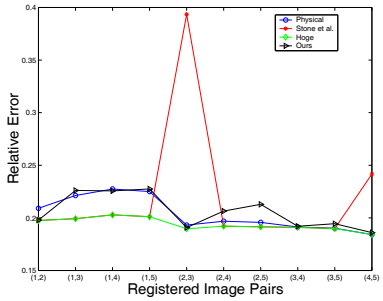


Fig. 6. The relative error in our registration compared to the physical calibration, Stone et al., and Hoge, for the MRI data set.

6. REFERENCES

[1] E. De-Castro and C. Morandi. Registration of translated and rotated images using finite Fourier transforms. *IEEE Trans. PAMI*, vol. 9(no. 5):700–703, 1987.

[2] P.A. Van den Elsen, E.J.D. Pol, and M.A. Viergever. Medical image matching-a review with classification. *IEEE Eng. Med. Biol. Mag.*, 12:pp. 26–39, 1993.

[3] H. Foroosh, J. Zerubia, and M. Berthod. Extension of phase correlation to sub-pixel registration. *IEEE Trans. Image Processing*, vol. 11(3):pp. 188–200, 2002.

[4] H. (Shekarforoush) Foroosh, M. Berthod, and J. Zerubia. Subpixel image registration by estimating the polyphase decomposition of cross power spectrum. In *Proc. CVPR*, pages pp. 532–537, 1996.

[5] W.S. Hoge. Subspace identification extension to the phase correlation method. *IEEE Trans. Medical Imaging*, 22(2):277–280, 2003.

[6] M. Irani and S. Peleg. Improving resolution by image registration. *Graphical Models and Image Processing*, 53(3):231–239, 1991.

[7] C. D. Kuglin and D. C. Hines. The phase correlation image alignment method. In *Proc. Int. Conf. on Cybernetics and Society*, pages 163–165, 1975.

[8] N. Nguyen, P. Milanfar, and G.H. Golub. A computationally efficient image superresolution algorithm. *IEEE Transactions on Image Processing*, 10(4):573–583, 2001.

[9] A. V. Oppenheim, R. W. Schafer, and with J.R. Buck. *Discrete-Time Signal Processing*. Prentice-Hall International Inc., 2nd edition, 1989.

[10] H. Shekarforoush and R. Chellappa. Data-driven multi-channel super-resolution with application to video data. *JOSA-A*, vol. 16:pp. 481–492, 1999.

[11] 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(10):pp. 2235–2243, 2001.

[12] A. Tikhonov and A. Arsenin. *Solutions of Ill-posed Problems*. Winston & Sons, 1977.

[13] A.H. Zemanian. *Distribution Theory and Transform Analysis*. Dover Publications Inc., 1965.