Available online at www.sciencedirect.com
ScienceDirect
Signal Processing 87 (2007) 124–133
SIGNAL PROCESSING
www.elsevier.com/locate/sigpro

A projection-based extension to phase correlation image alignment

Y. Keller[a,*], A. Averbuch[b]

[a]Department of Mathematics, Yale University, CT, USA
[b]School of Computer Science, Tel Aviv University, Tel Aviv, Israel

Received 30 September 2005; received in revised form 26 April 2006; accepted 28 April 2006
Available online 15 June 2006

Abstract

Phase-correlation (PC) is a computationally efficient scheme for estimating 2-D and 3-D translations. We present a masking operator that significantly improves the accuracy and robustness of the PC scheme. The masking projects the estimated correlation function on the space of correlation functions that result from a certain range of translations, while rejecting components that are unrelated to the estimated motion. Thus, the registration accuracy is improved by an order of magnitude, especially when registering noisy images and volumes. The scheme is also shown to improve the registration of rotated images in the Fourier domain.

Keywords: Fourier based; Phase correlation; Image registration; Robust alignment

1. Introduction

Image and volume registration play a vital role in applications such as video compression [1,2], video enhancement [3], scene representation [4,5], medical data processing [6,7] and range images registration [8] to name a few. Various computational techniques were suggested, such as gradient methods [1,4,5] and fast Fourier transform (FFT) based schemes [9,10]. Phase-correlation (PC) schemes [9,11–15] are able to estimate large rotations, scalings and translations, based on the phase shift property [16] of the Fourier transform.

Denote by $\widehat{I}(\vec{\omega}) = \mathfrak{F}\{I(\vec{t})\}$ the Fourier transform of I, where in the 2-D case $\vec{\omega} = (\omega_x, \omega_y)$, $\vec{t} = (t_x, t_y)$ and in the 3-D case $\vec{\omega} = (\omega_x, \omega_y, \omega_z)$ and $\vec{t} = (t_x, t_y, t_z)$.

Then

$$\mathfrak{F}\{I(\vec{t} + \vec{\Delta})\} = \widehat{I}(\vec{\omega})\, e^{j(\vec{\omega}\cdot\vec{\Delta})}, \tag{1}$$

where $\vec{\Delta} = (\Delta_x, \Delta_y)$ and $\vec{\Delta} = (\Delta_x, \Delta_y, \Delta_z)$ for the 2-D and 3-D cases, respectively.

Let the input images/volumes $I_1$and I_2 be related by

$$I_1(\vec{t} + \vec{\Delta}) = I_2(\vec{t}) \tag{2}$$

then

$$\widehat{I}_1(\vec{\omega})\, e^{j(\vec{\omega}\cdot\vec{\Delta})} = \widehat{I}_2(\vec{\omega}) \tag{3}$$

*Corresponding author. Tel.: +1 203 7777 982.
E-mail address: yosi.keller@yale.edu (Y. Keller).

0165-1684/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2006.04.013

and the phase difference matrix is given by

$$\widehat{C}(\vec{\omega}) = \frac{\widehat{I}_2(\vec{\omega})}{\widehat{I}_1(\vec{\omega})} = e^{j(\vec{\omega}\cdot\vec{\Delta})}. \quad (4)$$

The translation $\vec{\Delta}$ given by $\widehat{C}(\vec{\omega})$ can be recovered in either the spatial or frequency domains. The *spatial PC function* is given by [9,10,17]

$$C(\vec{t}) \triangleq \Im^{-1}\{\widehat{C}(\vec{\omega})\} = \delta(\vec{t} + \vec{\Delta}) \quad (5)$$

and $\vec{\Delta}$ is recovered by locating the maximum value of $C(\vec{t})$

$$\vec{\Delta} = \arg\max_{(\vec{t})}\{C(\vec{t})\}. \quad (6)$$

This scheme was proven to robustly estimate large translations where the corresponding overlap between the registered images is as small as 40% [13]. Due to its robustness it was also applied to rotation estimation [10] by aligning the magnitudes of the polar Fourier transforms of the input images. In such a representation, rotations are reduced to translations that can be estimated by the PC scheme.

Eq. (6) is computed using a fixed grid, hence, it is limited to integer pixel accuracy. $C(\vec{t})$ is a delta-like function for integer translations while non-integer translations cause the peak of $C(\vec{t})$ to spread across neighboring pixels, subsequently degrading the accuracy of the translation estimate. Furthermore, if the images are not spatially band-limited when sampled, aliasing will occur and Eq. (4) will not be valid for all frequencies [13].

The common approach for identifying non-integer shifts in the spatial domain is to apply bilinear, Lagrange, or other interpolation methods. But the PC function is not a smooth function, as it is the Dirichlet kernel [17] (see Section 2.1). Thus, its interpolation will not significantly improve the accuracy. Alternatively, one can work directly in the Fourier domain to identify the sub-pixel shifts. Such methods were described in [12,13,18], where linear regression was used to fit the phase values to a 2-D linear function

$$\vec{\omega}\cdot\vec{\Delta} = \arg\widehat{C}(\vec{\omega}). \quad (7)$$

These schemes provide sub-pixel accuracy as Eq. (4) holds for sub-pixel translations. Solving Eq. (7) directly using linear regression might prove inaccurate [17] due to aliasing and phase wrapping of the spectra at low-frequencies. Thus, an iterative solution to phase unwrapping was suggested in [12] and two approaches for modeling aliasing effects were given in [13]. Stone [13] also suggests masking the input images both in the spatial and frequency domains to improve the handling of the aliasing and noise effects.

The authors of [17] claim the regression to be inaccurate as it requires to fit a plane to the noisy phase data and provide an alternative extension to sub-pixel accuracy. First, an integer-precision alignment is obtained and then a sub-pixel refinement is derived by analyzing the shape of $C(\vec{t})$ around its maxima.

The work in [14,15] is conceptually related to the algorithm presented here. The method is based on the observation that a 'noise-free' PC function is of rank one. Hence, for a 'noisy' PC matrix, the sub-pixel motion estimation problem can be recast as finding the rank one approximation of that matrix, using the singular value decomposition (SVD). This improves the accuracy and robustness to noise without using interpolation.

In the context of the proposed scheme, the rank one approximation used in [14,15], is a projection operator, which projects $\widehat{C}(\vec{\omega})$ into the space of rank one linear functions. Other components related to noise, aliasing and phase wrapping are suppressed.

In order to compensate for intensity changes, Eq. (4) is often reformulated as

$$\widehat{C}_N(\vec{\omega}) \triangleq \frac{\widehat{I}_1^*(\vec{\omega})\widehat{I}_2(\vec{\omega})}{|\widehat{I}_1(\vec{\omega})||\widehat{I}_2(\vec{\omega})|} \quad (8)$$

and $\widehat{C}_N(\vec{\omega})$ replaces $C(\vec{\omega})$, where $\widehat{C}_N$ denotes the normalized PC [9] and $\widehat{I}_1^*$ is the complex conjugate of $\widehat{I}_1$.

In this paper we introduce a masking operator that projects the correlation function $C(\vec{t})$ into the space of correlation functions that result from a certain range of translations, thus attenuating the noise associated with the estimation of the phase-shifts by linear regression, especially when that noise is additive white Gaussian noise (AWGN). We show that the noise reduction of ordinary image sizes (256×256) is of the order of 10^3 and the registration accuracy is improved by an order of magnitude. We emphasize the applicability of the proposed scheme to sub-pixel registration and compare our results with those of [14,15,17]. These are considered the state of the art in PC-based sub-pixel registration. In particular, our approach is complementary to the regression based methods

[12,13], and to the PCSVD-based approach in [14,15]. Our scheme and [14,15] project $\widehat{C}(\vec{\omega})$ to different functional spaces, whose intersection is the space of linear phase exponential with limited spatial support.

Although the proposed approach improves the estimation of translations, it can also be applied to rotation estimation. Fourier-based schemes for rotation estimation [8,10,19,20] estimate rotations by reducing them to translations in a polar Fourier domain which are then estimated using PC. The polar Fourier representations are interpolated and are thus noisy. As the proposed scheme is especially suitable for registering noisy inputs, it can also improve the rotation estimation in [10] and we provide such experimental results in Section 3.

The rest of the paper is organized as follows: the masking operator is derived in Section 2 while the experimental results for 2-D and 3-D registration are given in Section 3.

2. The correlation masking operator

This section presents the masking operator. In Section 2.1 we show that the energy of the discrete PC function is concentrated in a small support. Thus, by zeroing it out beyond a limited support in Section 2.2, most of the noise is attenuated while retaining the components related to the motion. This makes the operator applicable to accurate sub-pixel registration.

2.1. Correlation function analysis

Consider the 1-D discrete phase correlation function

$$C(k) \triangleq \mathrm{e}^{\mathrm{j}k\Delta\omega\Delta x}, \quad 0 \leqslant k < N, \tag{9}$$

where N is the length of the signal, $\Delta\omega = 2\pi/N$ is the spacing of the phase axis and Δx is the shift. The outcome of Eq. (5) (the continuous case) is replaced by the inverse DFT of $C(k)$

$$\begin{aligned} C(n) &= \mathfrak{F}^{-1}\{C(k)\} \\ &= \frac{\sin(\pi(\Delta x + n))}{\sin(\pi(\Delta x + n)/N)} \\ &= \frac{\sin(\pi(\widetilde{\Delta x} + \lfloor \Delta x \rfloor + n))}{\sin(\pi(\widetilde{\Delta x} + \lfloor \Delta x \rfloor + n)/N)} \\ &= D(n - \Delta x), \end{aligned} \tag{10}$$

where $\Delta x = \widetilde{\Delta x} + \lfloor \Delta x \rfloor$, $\lfloor \Delta x \rfloor$ is the integer part of Δx and $D(n - \Delta x)$ is the Dirichlet kernel [16] shown in Fig. 1a. For integer pixel shifts ($\widetilde{\Delta x} = 0$), $D(n - \Delta x) = \Delta(n - \lfloor \Delta x \rfloor)$ (the Dirichlet delta function [16]) and for sub-pixel shifts it is split into several side-lobes. Most of its energy is concentrated in a limited spatial support around the integer part of the translation. An approximation of this support can be derived by considering the support needed to contain a certain percentage of the Dirichlet kernel's energy

$$\begin{aligned} \sum_{n=-R}^{R} \frac{\sin(\pi(\widetilde{\Delta x} + n))}{\sin(\pi(\widetilde{\Delta x} + n)/N)} &\approx \sum_{n=-R}^{R} \frac{\sin(\pi(\widetilde{\Delta x} + n))}{\pi(\widetilde{\Delta x} + n)/N} \\ &= N \sum_{n=-R}^{R} \frac{\sin(\pi(\widetilde{\Delta x} + n))}{\pi(\widetilde{\Delta x} + n)}. \end{aligned} \tag{11}$$

Denote

$$P(R) = \frac{N \sum_{n=-R}^{R} \dfrac{\sin(\pi(\widetilde{\Delta x} + n))}{\pi(\widetilde{\Delta x} + n)}}{N \sum_{n=-\infty}^{\infty} \dfrac{\sin(\pi(\widetilde{\Delta x} + n))}{\pi(\widetilde{\Delta x} + n)}}, \tag{12}$$

where $P(R)$ is the energy concentration of $C(n)$ in an interval $[-R, R]$ relative to the total energy. It follows that $P(R)$ is invariant to the length of the input signal N. Hence, it suffices to compute $P(R)$ once for various values of R. Such a computation is given in Figs. 1c and d for the 1-D and 2-D cases, respectively. We see that it suffices to use a support of $R = 5$, to retain 95% of the energy of $C(n)$. As expected, the largest support R should be used for $\widetilde{\Delta x} = 0.5$. For a given energy concentration $P(R)$ and a maximal sub-pixel motion $\widetilde{\Delta x}$, R can be computed numerically utilizing the fact that $P(R)$ is a monotonic increasing function and that R takes only integer values. When a maximal value of $\widetilde{\Delta x}$ cannot be assumed, we set $\widetilde{\Delta x} = 0.5$.

Due to the separability of the DFT in higher dimensions, $\widehat{C}(\vec{\omega})$ can be decomposed along different axes such that

$$\widehat{C}(\vec{\omega}) = \mathrm{e}^{\mathrm{j}(\vec{\omega} \cdot \vec{\Delta})} = \prod_{k=1}^{N} \mathrm{e}^{\mathrm{j}\omega_k \Delta x_k}, \tag{13}$$

$$C(\vec{t}) = \mathfrak{F}^{-1}\{\mathrm{e}^{\mathrm{j}(\vec{\omega} \cdot \vec{\Delta})}\} = \prod_{k=1}^{N} D(x_k) \tag{14}$$

and similar energy concentration properties are observed in the 2-D (Fig. 1b) and 3-D cases.

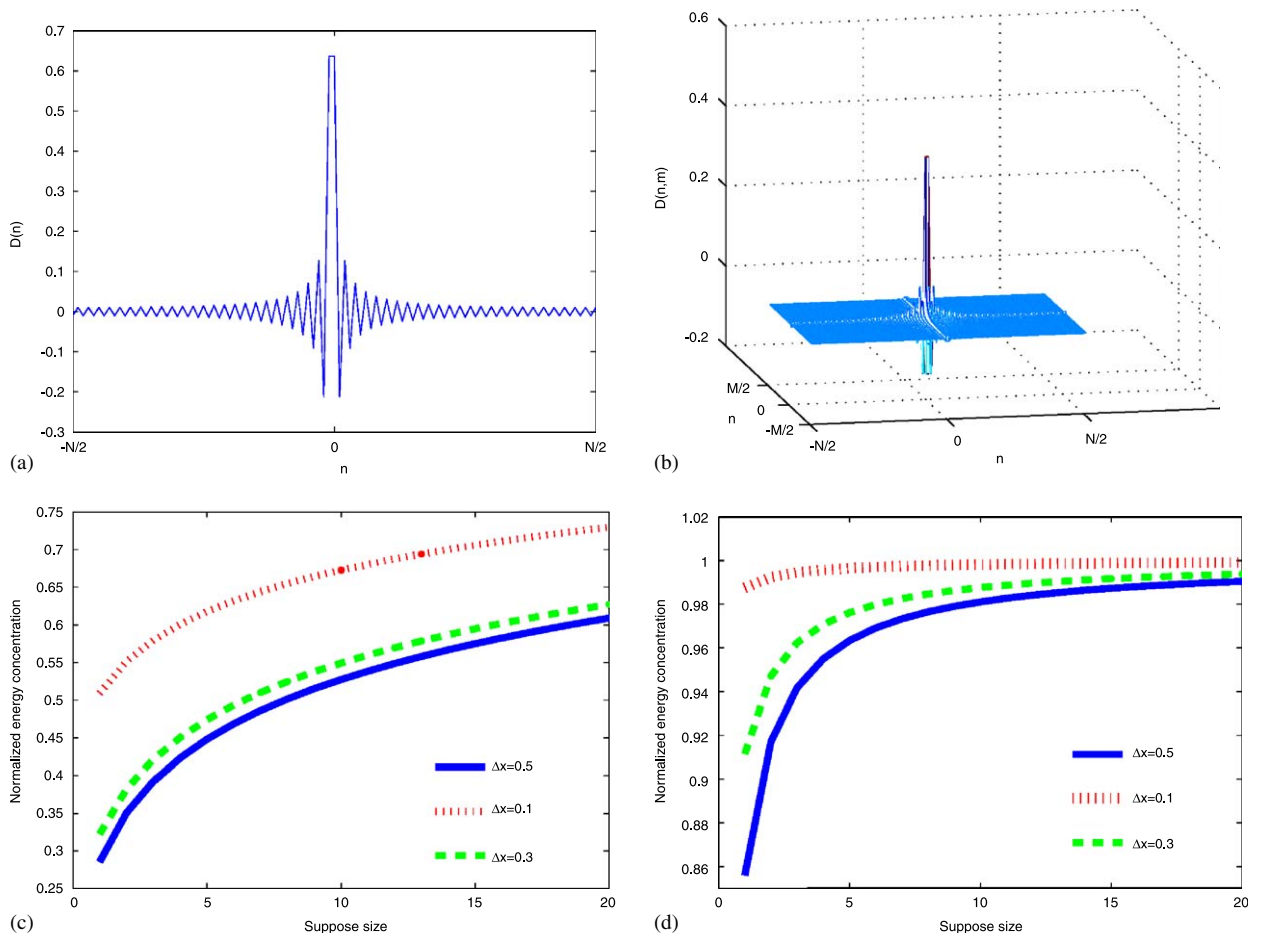


Fig. 1. The energy concentration property of the phase correlation function. In both the 1-D (a) and 2-D (b) cases, the energy of the Dirichlet kernel is concentrated in a small support around the integral part of the translation. (c) and (d) show the energy concentrations as a function of the support in the 1-D and 2-D cases, respectively.

2.2. The correlation function masking operator

The masking operator aims to improve the registration of noisy images, where the noise is assumed to be AWGN. The energy of WGN is spread evenly over the entire spatial domain [16], while the energy of the PC function $C(\vec{t})$ was shown (in the previous section) to be concentrated in a small spatial support. We build upon that to derive a masking operator that limits the support of $C(\vec{t})$ to the region where its energy is concentrated. Thus, given the input images I_1, I_2 we can compute the PC function $C(\vec{t})$,and restrict its spatial support. By taking into account the range of possible shifts D, given by

$$D = \{(\mathrm{d}x, \mathrm{d}y), \mathrm{d}x_{\min} \leqslant \mathrm{d}x \leqslant \mathrm{d}x_{\max}, \mathrm{d}y_{\min} \leqslant \mathrm{d}y \leqslant \mathrm{d}y_{\max}\}$$

and

$$(\delta x, \delta y) = (\mathrm{d}x_{\max} - \mathrm{d}x_{\min}, \mathrm{d}x_{\max} - \mathrm{d}x_{\min}),$$

the support of $C(\vec{t})$ is then restricted to

$$\widehat{D} = \left\{(\mathrm{d}x, \mathrm{d}y), \mathrm{d}x_{\min} - \frac{R}{2} \leqslant \mathrm{d}x \leqslant \mathrm{d}x_{\max} + \frac{R}{2}, \mathrm{d}y_{\min} - \frac{R}{2} \leqslant \mathrm{d}y \leqslant \mathrm{d}y_{\max} + \frac{R}{2}\right\}.$$

This is exemplified in Fig. 2.

The noise energy is significantly attenuated, while the energy of $C(\vec{t})$ is reduced slightly. The support size R is computed by Eq. (12) and Fig. 1 such that 90% of the energy of $C(\vec{t})$ is retained. We note that if the registration is bootstrapped by computing the

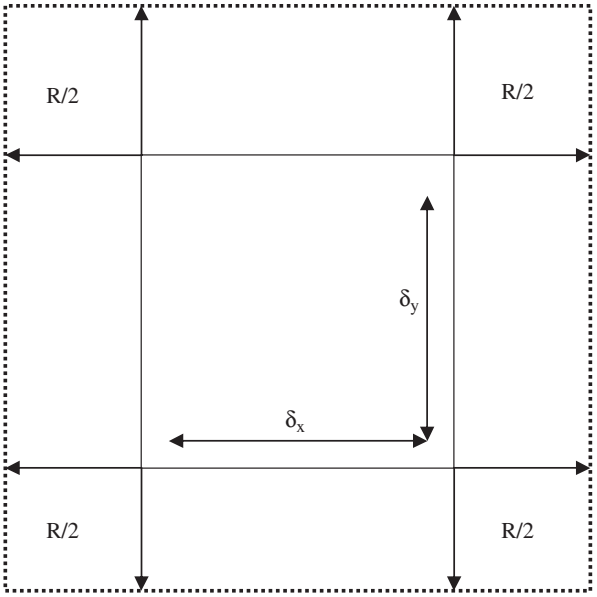


Fig. 2. The proposed phase-correlation masking operator. The masking operator restricts the support of the spatial correlation function. The inner rectangle $[-\delta x/2, \delta x/2] \times [-\delta y/2, \delta y/2]$ is the interval of uncertainty of the translations and the phase-correlation peak might be anywhere within this rectangle. The dotted line marks the extension of the uncertainty support by R. The masking operator zeros the correlation function beyond this support.

integer part of the shift first, we have $(\delta x, \delta y) = (0, 0)$ and the gain of the masking operator is maximized.

The attenuation of the noise energy is given by the ratio of the areas

$$\Delta\text{WGN} \approx \frac{N \cdot M}{(R+\delta x)(R+\delta y)}, \tag{15}$$

where N and M are the initial sizes of the image. The overall improvement is given by

$$\Delta\text{SNR} \approx \frac{N \cdot M \cdot P}{(R+\delta x)(R+\delta y)}, \tag{16}$$

where P is the percentage of the energy of the phase correlation function $C(\vec{t})$, which is retained. For example, for $N = M = 512$, $R = 7$, $(\delta x, \delta y) = (10, 10)$ (a typical effective size used in our experiments) we get an improvement of

$$\Delta\text{SNR} \approx \frac{512^2 \cdot 0.9}{(7+10)(7+10)} = 816 = 29\,\text{dB}. \tag{17}$$

For a similar 3-D scenario, the gain is the ratio of the volumes

$$\Delta\text{SNR} \approx \frac{512^3 \cdot 0.9}{(7+10)^3} = 2.45 \cdot 10^4 = 43\,\text{dB}. \tag{18}$$

The smaller the translation uncertainties δx and δy, the better the SNR improvement. Moreover, the SNR improvement is unrelated to the spectral content of the registered images. It is due to the attenuation of the noise, as its component outside of the mask is zeroed, while the signal is known to be concentrated within the mask. Thus, Eq. (16) holds for any pair of registered images and volumes. Fig. 3 exemplifies the denoising of a phase correlation function using the proposed masking operator.

The phase of the masked PC function (shown in Fig. 3d) is a scaled replica of the original phase (shown in Fig. 3c). Let $(\mathrm{d}x, \mathrm{d}y)$ be the pixelwise shift in the input images, then $\Delta\varphi$, the phase shift of the original images is given by

$$(\Delta\varphi_x, \Delta\varphi_y) = \left(2\pi\frac{\mathrm{d}x}{M}, 2\pi\frac{\mathrm{d}y}{N}\right). \tag{19}$$

By projecting the PC function, its dimensions are reduced, but the initial pixelwise shifts are retained. Hence, $(\widehat{\Delta\varphi_x}, \widehat{\Delta\varphi_y})$ the phase shift of the masked PC function (of dimensions $(R+\delta x) \times (R+\delta y)$) is given by

$$(\Delta\varphi_x, \Delta\varphi_y) = \left(2\pi\frac{\mathrm{d}x}{R+\delta x}, 2\pi\frac{\mathrm{d}y}{R+\delta y}\right). \tag{20}$$

Thus, in order to recover the original phase shift $(\Delta\varphi_x, \Delta\varphi_y)$, the phase recovered by Fourier domain schemes such as [14,15], has to be divided by a factor of $(M/(R+\delta x), N/(R+\delta y))$. As we use $(R+\delta x) \ll M$ and $(R+\delta y) \ll N$, the phase shift in Eq. (20) is significantly larger than the original phase shift in Eq. (19). When visualized (similar to Fig. 3d) we get a higher frequency signal compared to Fig. 1c in [15]. The application of the phase masking operator to motion estimation is summarized in Algorithm 1.

Algorithm 1. Image and volume registration using the masking operator

1: *Input*: Two input images or volumes I_1 and I_2.
2: Apply Eq. (8) to compute the normalized phase correlation function $\widehat{C}_N(\vec{\omega})$, while masking the input images as in [13].
3: Compute the spatial correlation function $C_N(\vec{t}) = \Im^{-1}\{\widehat{C}_N(\vec{\omega})\}$.

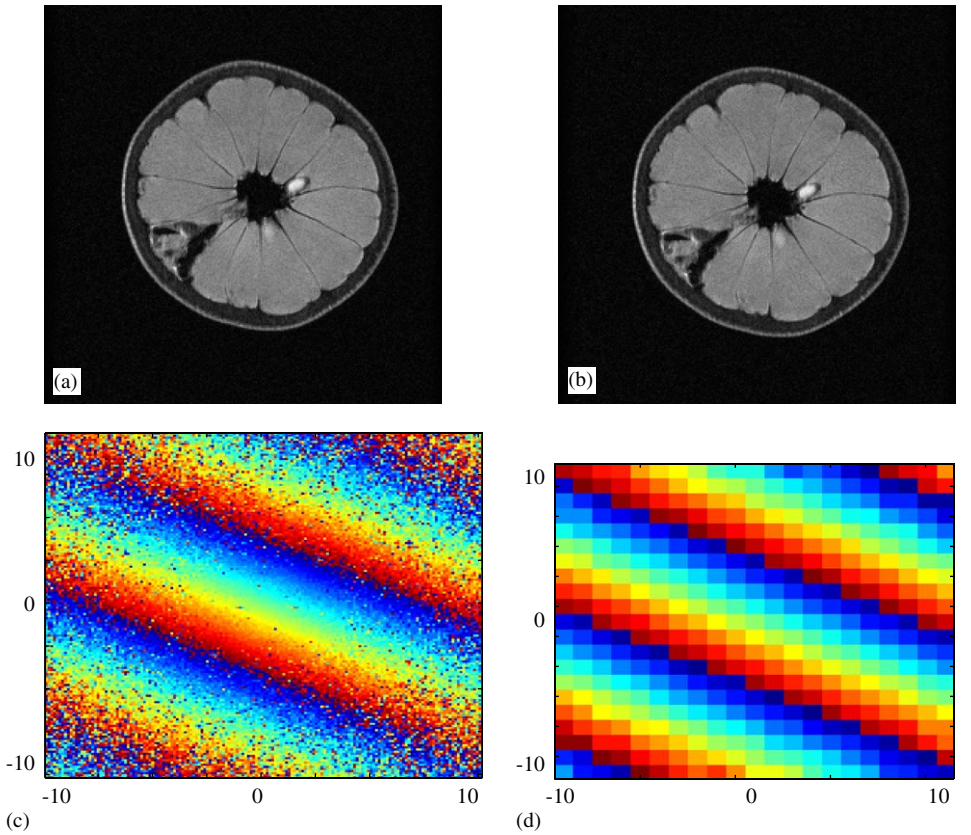


Fig. 3. An example of the phase denoising achieved by the masking operator. The operator is applied to the phase-correlation function of the Grapefruit MRI images (a) and (b) (used by courtesy of W.S. Hoge). The images are related by a translation of $(\delta x, \delta y) = (10, 10)$, and are of size 256^2. By masking the phase correlation function (depicted in (c)) using $R = 11$ the phase in (d) is denoised.

4: Apply the masking operator to $C_N(\overrightarrow{t})$ and compute the projected phase correlation function $C_N^P(\overrightarrow{t})$

$$C_N^P(\overrightarrow{t}) = \begin{cases} C_N(\overrightarrow{t}), & \begin{array}{l} \mathrm{d}x_{\min} - \dfrac{R}{2} \leqslant \mathrm{d}x \leqslant \mathrm{d}x_{\max} + \dfrac{R}{2}, \\ \mathrm{d}y_{\min} - \dfrac{R}{2} \leqslant \mathrm{d}y \leqslant \mathrm{d}y_{\max} + \dfrac{R}{2}, \end{array} \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

in the 2-D case, or

$$C_N^P(\overrightarrow{t}) = \begin{cases} C_N(\overrightarrow{t}), & \begin{array}{l} \mathrm{d}x_{\min} - \dfrac{R}{2} \leqslant \mathrm{d}x \leqslant \mathrm{d}x_{\max} + \dfrac{R}{2}, \\ \mathrm{d}y_{\min} - \dfrac{R}{2} \leqslant \mathrm{d}y \leqslant \mathrm{d}y_{\max} + \dfrac{R}{2}, \\ \mathrm{d}z_{\min} - \dfrac{R}{2} \leqslant \mathrm{d}z \leqslant \mathrm{d}z_{\max} + \dfrac{R}{2}, \end{array} \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

for the 3-D case.

5: Compute the Fourier transform of the masked correlation function $C_N^P(\overrightarrow{\omega}) = \mathfrak{F}\{C_N^P(\overrightarrow{t})\}$.
6: Apply a motion estimation scheme such as the PCSVD [13–15] to the projected correlation function $C_N^P(\overrightarrow{\omega})$.

3. Experimental results

The proposed scheme was experimentally verified by applying the masking operator to translation and rotation estimation. First, we align sub-pixel shifted replica of the images given in Fig. 4 and the volumes shown in Fig. 7. Then we applied the masking operator to the rotation estimation scheme given in [10]. We show that in all three cases the registration accuracy is improved.

In order to measure the precision of the registration algorithm, we prepared pairs of images and volumes, that were related by known sub-pixel shifts. These would act as ground truth for the

Fig. 4. The images used for verification: (a) Grapefruit MRI image (courtesy of W.S. Hoge). (b) CT slice image. (c) MRI slice image.

experiments. For that we applied the scheme suggested by Stone in [13] and later used by Foroosh in [17]. The idea is to start with a single high-resolution image and create two partial replicas of it, which are shifted by integer amounts. This can be easily done by cropping. Then, by filtering and downsampling each of the replicas, the relative shifts become fractional. This procedure is summarized in Algorithm 2.

Algorithm 2. Preparing the ground truth images and volumes

1: Given a pair of input images or volumes I_1 and I_2.

2: Let the sub-pixel motion one would like to induce be: $\vec{\Delta} = (\vec{k}/\vec{m})$ where $(\Delta x, \Delta y) = (k_1/m_1, k_2/m_2)$ or $(\Delta x, \Delta y, \Delta z) = (k_1/m_1, k_2/m_2, k_3/m3)$, for the 2-D and 3-D cases, respectively.

3: I_1 and I_2 are smoothed using a Gaussian kernel of size $\max(\vec{m})$.

4: The overlapping supports of $I_1(\vec{t})$ and $I_2(\vec{t} + \vec{k})$ are extracted.

5: The extracted images are downsampled by factors of $\vec{m}$ to generate the sub-pixel shifted images.

The registration of the images was compared with three methods: the correlation function shape analysis scheme given in [17], SVD decomposition of the phase correlation $\widehat{C}_N(\vec{\omega})$ (PCSVD) [15] and a combination of the proposed masking operator together with the SVD decomposition, denoted as the projected PCSVD (PPCSVD). For that we applied a masking operator with a radius of 11 before computing SVD of $\widehat{C}_N(\vec{\omega})$. This corresponds to retaining 90% of the initial energy of the correlation function. We did not consider [13], since it was shown in [15] to be less accurate than the PCSVD. Since the algorithm in [17], is unable to estimate shifts greater than one pixel, we first computed the integer part of the translation using phase correlation (Eq. (6)).

Table 1 shows the performance of the registration on various images. It follows that the PPCSVD was consistently more accurate than the PCSVD by a factor of 2–3. Both were able to achieve an accuracy of $10^{-2}-10^{-3}$ and outperform the accuracy of [17]. This is consistent with expectations, given that [17] uses no filtering of $\widehat{C}_N$. Fig. 5 depicts the noise filtering properties of the PPCSVD, where we present the noisy MRI image and its noisy and projected phase. Given the phase correlation function, we found the PCSVD to be the most effective scheme for computing the motion. It is evident from Fig. 5 that it is better to apply the PCSVD to Fig. 5c rather than to Fig. 5b.

This result was further verified by systematically registering sets of noisy images and volumes. A test pair with a sub-pixel shift of $(\Delta x, \Delta y) = (0.5, 0.5)$ of the images in Fig. 4 was generated. White Gaussian noise was added to the images, which were then registered using a masking operator with a radius of $R = 11$, where the form of the masking rectangle is depicted in Fig. 2. In order to analyze the registration of noisy images we tested for 50 different standard deviations of the noise $n(x, y) \sim N(\sigma = 4 - 200, \mu = 0)$. Each experiment was repeated 100 times for statistical validity. Overall, for each noise level 300 experiments were

conducted. The alignment error measure ε is given by $\varepsilon = \sqrt{\varepsilon_x^2 + \varepsilon_y^2}$, where ε_x and ε_y are the registration errors in the x- and y-axes, respectively. The results, given in Fig. 6, demonstrate the improvement achieved by the PPCSVD and the denoising properties of the proposed masking operator.

Similar simulations were conducted for 3-D registration of the volumes given in Fig. 7. For each volume, we shifted the volumes by $(\Delta x, \Delta y, \Delta z) = (0.5, 0.5, 0.5)$ and added a 3-D WGN$\sim(\sigma = 4 - 200, \mu = 0)$. Each simulation was repeated 100 times and the errors in all three axes were combined by computing the Euclidean norm of the error: $\varepsilon = \sqrt{\varepsilon_x^2 + \varepsilon_y^2 + \varepsilon_z^2}$. The results, shown in Fig. 7c, are in accordance with the 2-D results: the masking operator significantly improves the registration of noisy data.

Next, we applied our approach to rotation estimation by registering the images in Fig. 8 using the state-of-the-art scheme given in [10]. Each image was registered to a replica which was randomly rotated in the range of $[-20°, 20°]$ and translated in the range of $[-10, 10]$ pixels. The magnitudes of the

Table 1
Registration results for noise-free medical images

	Accurate		SVD		Foroosh [17]		PPCSVD	
	Δx	Δy	ε_x	ε_y	ε_x	ε_y	ε_x	ε_y
Grapefruit MRI	2.50	−2.50	200	500	400	800	20	70
	0.25	1.50	20	1000	20	300	0.1	2
	−1.25	−4.50	60	100	200	400	2	4
CT slice	0.167	−0.50	20	40	70	30	2	2
	0.5	3.25	20	3	30	8	2	1
	−1.33	−1.167	30	4	100	30	3	4
MRI slice	−2.33	−2.83	10	70	5	90	3	3
	−1.33	1.33	50	4	70	20	10	2
	−1.66	−1.66	9	10	30	70	5	10

The shifts are in terms of pixels and errors are in terms of 1/1000 of a pixel. The true shift values $(\Delta x, \Delta y)$ are given in the left column (accurate). Each pair of columns shows the horizontal and vertical registration error $(\varepsilon_x, \varepsilon_y)$. The best accuracy is achieved by applying the proposed masking operator together with the SVD-based scheme (third row). It is compared to the results of the PCSVD scheme given in [15] and Foroosh's method [17].

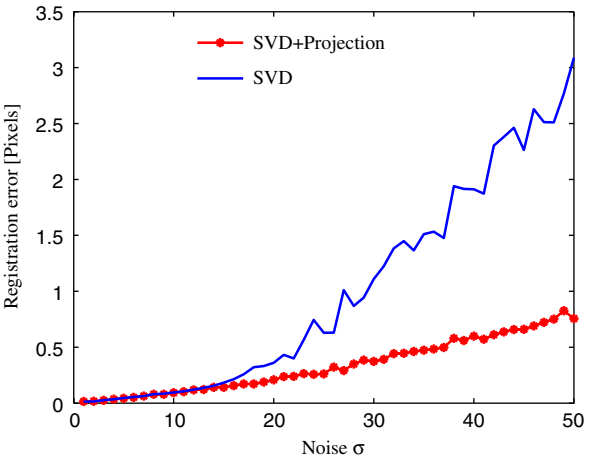
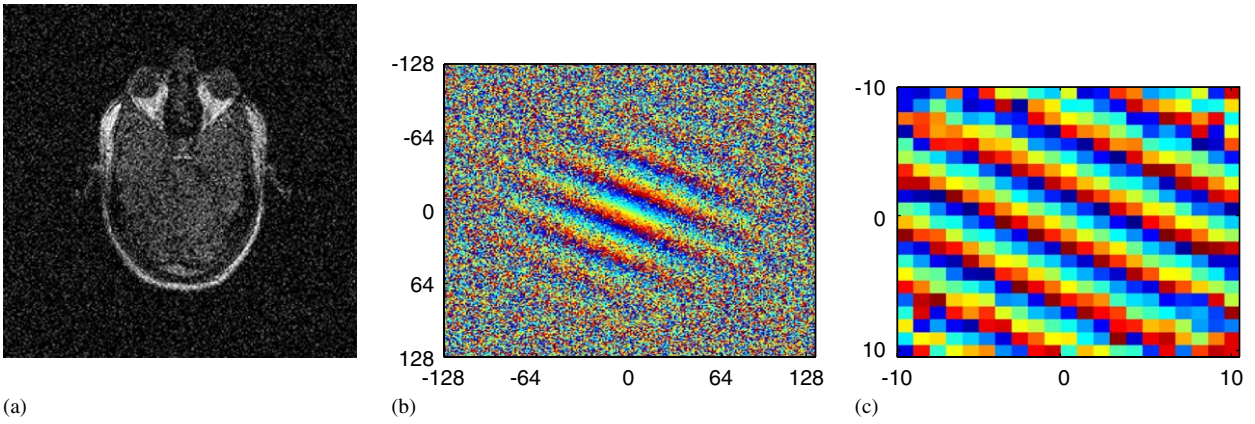


Fig. 6. Average accuracy of 2-D registrations with added noise levels. The denoising properties of the masking operator are demonstrated by the improvement in the registration accuracy for high noise levels. The combined error ε was computed by the Euclidean norm of the errors $\varepsilon = \sqrt{\varepsilon_x^2 + \varepsilon_y^2}$.



(a) (b) (c)

Fig. 5. An example of the noise attenuation achieved by the masking operator when applied to noisy images. (a) The input noisy image ($\sigma = 40$) was registered to a noisy replica translated by $(t_x, t_y) = (0.5, 0.5)$. (b) The phase of the noisy phase correlation function. (c) The phase of the projected phase correlation function. The lines of constant phase correspond to frequencies where $\omega_x t_x + \omega_y t_y = C$.

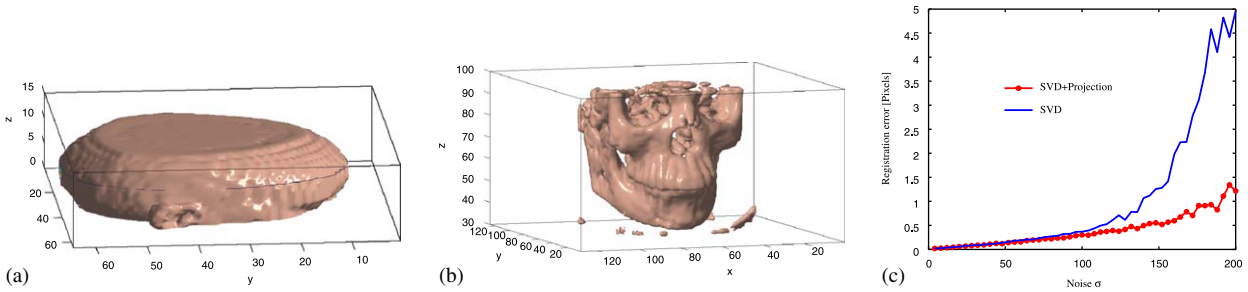


Fig. 7. Average accuracy of 3-D registration with varying added noise levels. (a) MRI scan of a human head. (b) CT scan of a human skull. (c) Applying the masking operator to the HOSVD-based scheme significantly improved the registration accuracy of the noisy volumes. The combined error ε was computed by the Euclidean norm of the errors $\varepsilon = \sqrt{\varepsilon_x^2 + \varepsilon_y^2 + \varepsilon_z^2}$.

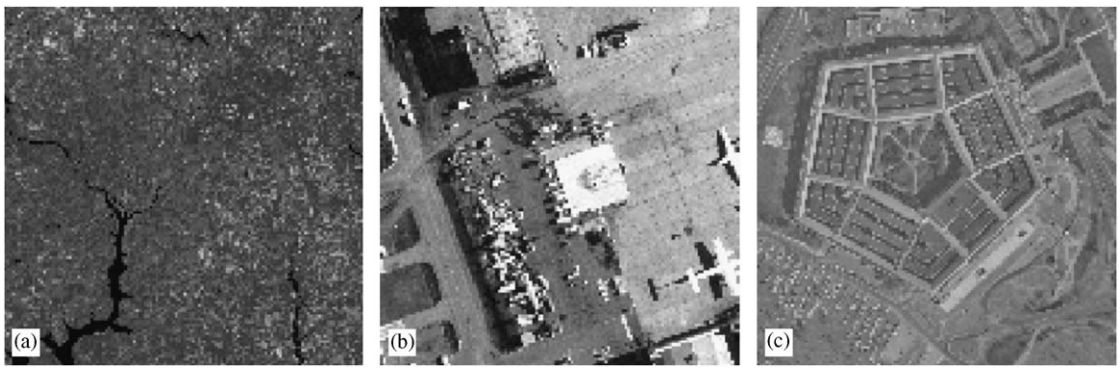


Fig. 8. The images used for applying the proposed approach to rotation estimation: (a) SAR image. (b) Airfield image. (c) Pentagon image.

Table 2
Rotation estimation results

	True rotation (deg.)	Phase correlation (deg.)	PPCSVD (deg.)
SAR	3.5	2.3	3.0
	15.7	16.2	16.1
	−17.6	−17.2	−17.5
Airfield	−5.8	−5.3	−5.4
	12.5	14.0	13.1
	−19.6	−18.7	−19.5
Pentagon	−14.4	−12.5	−13.9
	−11.8	−10.9	−11.6
	−12.0	−11.2	−11.4

The images were registered by the state-of-the-art scheme given in [10]. The rotations was reduced to translations of the Fourier magnitudes, which were estimated by phase correlation (middle column) and the proposed scheme (right column).

input images in the polar Fourier domain were computed by bilinear interpolation. The results given in Table 2 show that using the masking operator reduced the registration error. The standard deviation of the angular error of [10] was $\sigma = 0.560^\circ$ while the error achieved using the proposed scheme was $\sigma = 0.209^\circ$.

We attribute the improvement in the registration precision to the polar Fourier coefficients being computed by interpolation in the Fourier magnitude domain. Hence, the coefficients are noisy and oscillatory, making their registration especially suitable for the proposed algorithm, which excels in the registration of noisy data. These results are thus comparable to the registration results of the noisy images and volumes depicted in Figs. 5–7. In this work we advocate the use of the rectangular windows, as we also experimented with Gaussian and Hamming windows, where we recorded similar accuracies.

4. Summary and conclusions

This work presents an extension to the state-of-the-art phase correlation (PC) technique for image and volume registration. We introduce a masking

operator that is applied to the phase correlation function. Using the proposed operator improves the registration precision, especially for noisy data, where the noise is AWGN.

The operator projects the PC function to the space of functions related to a particular range of translations, while attenuating other phase components that are unrelated to the estimated motion. Thus, it better handles aliasing, phase warping and noise, making it suitable for the registration of medical images that are often noisy.

While the approach is complementary to other sub-pixel image registration approaches, we found the rank one PC scheme (PCSVD) given in [14,15] to be the most appropriate. When combined with the proposed method, a significant accuracy improvement was achieved over the PCSVD and the state-of-the-art method in [17]. It was also shown to improve the estimation of rotations in the Fourier domain.

In future we intend to apply the operator to the estimation of local motions based on PC. In such schemes [21], the motion is localized by computing the phase of a local Fourier transform of the input images.

References

[1] F. Dufaux, J. Konrad, Efficient, robust, and fast global motion estimation for video coding, IEEE Trans. Image Process. 9 (3) (2000) 497–501.

[2] A. Tekalp, Digital Video Processing, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[3] M. Irani, S. Peleg, Motion analysis for image enhancement: resolution, occlusion and transparency, J. Visual Comm. Image Representation 4 (4) (1993) 324–335.

[4] S. Mann, R. Picard, Virtual bellows: constructing high quality stills from video, IEEE International Conference Image Processing, Austin, TX, 1994, pp. 363–367.

[5] R. Szeliski, Image mosaicking for tele-reality applications, in: Proceedings of IEEE Workshop on Applications of Computer Vision, 1994, pp. 44–53.

[6] J. Maintz, M. Viergever, A survey of medical image registration, Medical Image Anal. 2 (1) (1998) 1–36.

[7] J. Pluim, J. Maintz, M. Viergever, Mutual information based registration of medical images: a survey, IEEE Trans. Medical Imaging 22 (8) (2003) 986–1004.

[8] L. Lucchese, G. Doretto, G. Cortelazzo, A frequency domain technique for range data registration, IEEE Trans. Pattern Anal. Mach. Intell. 24 (11) (2002) 1468–1484.

[9] C.D. Kuglin, D.C. Hines, The phase correlation image alignment method, IEEE Conf. Cybernet. Soc. 1 (1975) 163–165.

[10] S. Reddy, B.N. Chatterji, An FFT-based technique for translation, rotation, and scale-invariant image registration, IEEE Trans. Image Process. 3 (8) (1996) 1266–1270.

[11] P. Milanfar, Projection-based, frequency-domain estimation of superimposed translational motions, J. Opt. Soc. Amer. A 13 (11) (1996) 2151–2162.

[12] Y. Chou, H. Hang, A new motion estimation method using frequency components, J. Visual Comm. Image Representation 8 (1) (1997) 83–96.

[13] H. Stone, M. Orchard, E.-C. Chang, S. Martucci, A fast direct Fourier-based algorithm for subpixel registration of images, IEEE Trans. Geosci. Remote Sensing 39 (10) (2001) 2235–2243.

[14] W.S. Hoge, C.-F. Westin, Identification of translational displacements between *N*-dimensional data sets using the high order SVD and phase correlation, IEEE Trans. Image Process. 14 (7) (2005) 884–889.

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

[16] B. Porat, A Course in Digital Signal Processing, Wiley, New York, 1997.

[17] H. Foroosh, J. Zerubia, M. Berthod, Extension of phase correlation to subpixel registration, IEEE Trans. Image Process. 11 (3) (2002) 188–200.

[18] D.J. Fleet, Disparity from local weighted phase-correlation, IEEE Internat. Conf. Systems Man Cybernet. 1 (1994) 48–56.

[19] P. Milanfar, Two-dimensional matched filtering for motion estimation, IEEE Trans. Image Process. 8 (3) (1999) 438–443.

[20] Q. Chen, M. Defrise, F. Deconinck, Symmetric phase-only matched filtering of Fourier-mellin transforms for image registration and recognition, IEEE Trans. Pattern Anal. Mach. Intell. 6 (12) (1994) 1156–1168.

[21] H. Foroosh, M. Balci, Sub-pixel registration and estimation of local shifts directly in the Fourier domain, in: IEEE International Conference on Image Processing, ICIP 2004, 2004.