# LETTER High-Accuracy Sub-Pixel Registration for Noisy Images Based on Phase Correlation

Bei HE<sup>†a)</sup>, Student Member, Guijin WANG<sup>†b)</sup>, Member, Xinggang LIN<sup>†</sup>, Nonmember, Chenbo SHI<sup>†</sup>, Student Member, and Chunxiao LIU<sup>†</sup>, Nonmember

**SUMMARY** This paper proposes a high-accuracy sub-pixel registration framework based on phase correlation for noisy images. First we introduce a denoising module, where the edge-preserving filter is adopted. This strategy not only filters off the noise but also preserves most of the original image signal. A confidence-weighted optimization module is then proposed to fit the linear phase plane discriminately and to achieve subpixel shifts. Experiments demonstrate the effectiveness of the combination of our modules and improvements of the accuracy and robustness against noise compared to other sub-pixel phase correlation methods in the Fourier domain.

**key words:** phase correlation, sub-pixel, edge-preserving filter, confidenceweighted optimization

# 1. Introduction

Sub-pixel image registration plays a crucial role in computer vision and image processing, such as super resolution, denoising and medical reconstruction [1], [2]. Phase correlation has proved to be an attractive solution [4] among a wide variety of registration techniques.

Shift-estimation based on phase correlation falls into two categories: methods in the spatial and the Fourier domains. The methods in the spatial domain obtain the peak of the phase correlation surface as the shifts [3], [5], [7]. Foroosh [5] calculated the shifts by linear-weighting the maximum and sub-maximum peaks. J Ren [7] took into consideration the difference between two side neighbors of the maximum peak to improve the robustness against noise.

Methods in the Fourier domain fit the linear phase plane to achieve sub-pixel estimations [4], [8], [9]. Patrick [8] picked data points in the low-frequency band only of the linear phase plane to alleviate the effects of the high-frequency outliers. Keller [9] utilized the energy concentration property of the phase correlation surface and used the low-pass filter to suppress the noise. However, these methods utilize all picked data points in the linear phase plane equally when outliers of this plane, sometimes existing in the low-frequency band, might corrupt the estimations.

In this paper, a high-accuracy sub-pixel registration framework is proposed based on phase correlation in the

Manuscript received May 16, 2011.

Manuscript revised August 20, 2011.

a) E-mail: b-he08@mails.tsinghua.edu.cn

b) E-mail: wangguijin@tsinghua.edu.cn

DOI: 10.1587/transinf.E94.D.2541

Fourier domain. First, a denoising module is introduced, where the edge-preserving filter is proposed and utilized before the traditional phase correlation module to retain a more original image signal. Second, we propose a confidenceweighted optimization module to fit the linear phase plane discriminately which leads to accurate and robust estimations.

The rest of the paper is organized as follows. The formulation of phase correlation is introduced in Sect. 2. Our proposed scheme is presented in Sect. 3. Evaluation of our framework is given in Sect. 4, followed by a brief conclusion in Sect. 5.

#### 2. Phase Correlation Method

Suppose images  $I_1(\mathbf{x})$  and  $I_2(\mathbf{x})$  satisfy  $I_2(\mathbf{x}) = I_1(\mathbf{x} + \Delta \mathbf{x})$ where  $\mathbf{x} = (x, y)$  and  $\Delta \mathbf{x} = (\Delta x, \Delta y)$ . Their Fourier transforms are denoted by  $F_1(\mathbf{u})$  and  $F_2(\mathbf{u})$  where  $\mathbf{u} = (u, v)$ . According to the Fourier shift property, the **phase correlation** function  $C(\mathbf{u})$  is calculated as

$$C(\mathbf{u}) = \frac{F_2(\mathbf{u})}{F_1(\mathbf{u})} = \exp\{2\pi j \cdot \mathbf{u} \Delta \mathbf{x}\}.$$
 (1)

If we apply the inverse Fourier transform to  $C(\mathbf{u})$ , the **phase correlation surface**  $c(\mathbf{x})$  is obtained in Eq. (2). By locating the peak of the surface, we get the integer shifts.

$$c(\mathbf{x}) = \mathfrak{F}^{-1}\{C(\mathbf{u})\} = \delta(\mathbf{x} + \Delta \mathbf{x}).$$
(2)

The sub-pixel shifts can be estimated by fitting the **linear phase plane**  $A(\mathbf{u})$  calculated in Eq. (3).

$$A(\mathbf{u}) = \arctan(C(\mathbf{u})) = \mathbf{u}\Delta\mathbf{x}.$$
(3)

However, noise can not be neglected in real images and  $C(\mathbf{u})$  in Eq. (1) should be changed to  $C'(\mathbf{u})$ ,

$$C'(\mathbf{u}) = \frac{C(\mathbf{u}) + N_2(\mathbf{u})/F_1(\mathbf{u})}{1 + N_1(\mathbf{u})/F_1(\mathbf{u})},$$
(4)

where  $N_1(\mathbf{u})$  and  $N_2(\mathbf{u})$  are the Fourier transform of the noises in  $I_1(\mathbf{x})$  and  $I_2(\mathbf{x})$  respectively. When the Signal to Noise Ratio (SNR) is low, the phase correlation surface is turned to be unreliable and there would be many outliers existing in the linear phase plane. As a result, fitting all data points equally would degrade the estimation accuracy.

<sup>&</sup>lt;sup>†</sup>The authors are with the Dept. of Electronic Engineering, Tsinghua University, ROC.



Fig. 1 The global flowchart of our framework.

#### 3. Proposed Framework

Figure 1 shows the flowchart of our proposed framework. Before the step of phase correlation, we apply the denoising module with an edge-preserving filter. Then we divide the shift estimation into two steps: the coarse step and the refine one. In the coarse step, the integer shifts  $\Delta \mathbf{x}_1 = (\Delta x_1, \Delta y_1)$  are obtained by determining the peak of the phase correlation surface. After making compensation for the integer shifts, we achieve the decimal shifts  $\Delta \mathbf{x}_2 = (\Delta x_2, \Delta y_2)$  in the refine step. Finally, the integer and decimal parts are added as the ultimate shifts  $\Delta \mathbf{x}$ . Our contributions concentrate on the design of the edge-preserving denoising and the confidence-weighted optimization modules.

#### 3.1 Edge-Preserving Denoising

According to Eq. (4), noise existing in the original image would affect the fitting step. Especially when the SNR is low, the influence will even be amplified. Therefore, we denoise the images before phase correlation. As the noise mainly concentrates on the high-frequency band, the lowpass filter can alleviate the noise. However, lacking the capacity to separate the image signal from the noise, this type of the filter also reduces the high-frequency image signal. Therefore, fewer data points are available for plane fitting and would lead to less accurate estimations.

This paper introduces the edge-preserving filter [10] for noisy images, which is formulated as

$$I_{o}(\mathbf{x}) = \frac{1}{S} \sum_{\mathbf{m} \subseteq R(\mathbf{x})} I_{i}(\mathbf{m})$$
$$G_{\alpha}(||\mathbf{m} - \mathbf{x}||)G_{\beta}(|I_{i}(\mathbf{m}) - I_{i}(\mathbf{x})|),$$
(5)

where  $I_i$  and  $I_o$  are input and output images, R represents the neighbor area of the point **x**.  $G_\alpha(||\mathbf{m} - \mathbf{x}||)$  refers to the spatial Gaussian kernel with the variance  $\alpha$ , and  $G_\beta(|I_i(\mathbf{m}) - I_i(\mathbf{x})|)$  refers to the range Gaussian kernel with the variance  $\beta$ . These two kernels constitute our edge-preserving filter. The normalized item S is denoted as

$$S = \sum_{\mathbf{m} \subseteq R(\mathbf{x})} G_{\alpha}(\|\mathbf{m} - \mathbf{x}\|) G_{\beta}(|I_i(\mathbf{m}) - I_i(\mathbf{x})|).$$
(6)

In the flat region where the pixels have the similar range (corresponding to the low-frequency image signal), the edge-preserving filter acts like a low-pass filter. While in the edge region where the pixels range varies abruptly (corresponding to the high-frequency image signal), this filter can preserve more image signal by using the range Gaussian kernel. Consequently, in the low SNR, our scheme can still collect most of the original image signal and lead to more convincing estimations. Though the filter may change some complex edge locations, we preserve a majority of the image signal. This module is theoretically applicable as there is always noise existing in the real images and the experiments in Sect. 4 support that.

# 3.2 Confidence-Weighted Optimization

In the previous steps, we estimate and compensate for the integer shifts. In the refine step, the outliers in the linear phase plane caused by the noise and aliasing would corrupt the estimations of direct fitting. So we evaluate each data point with a confidence-weight and propose the weighted optimization to calculate the decimal shift. This implies the plane is fitted discriminately. The optimization criterion is defined as

$$\underset{\Delta \mathbf{x}_{2}}{\arg\min} \sum_{\mathbf{u}} \omega(\mathbf{u}) (\mathbf{u} \Delta \mathbf{x}_{2} - A(\mathbf{u}))^{2}, \tag{7}$$

where **u** represents the coordinate point (u, v),  $\omega(\mathbf{u})$  refers to the corresponding confidence-weight, and  $A(\mathbf{u})$  refers to the linear phase plane in Eq. (3). Accurate sub-pixel shifts can be resolved as a closed form solution defined as,

$$\Delta \mathbf{x}_2 = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{b},\tag{8}$$

where  $\mathbf{H} = (\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_M)^T$ ,  $\mathbf{H}_k = \sqrt{\omega(\mathbf{u})}\mathbf{u}^T$ ,  $\mathbf{b} = (b_1, b_2, \dots, b_M)^T$ ,  $b_k = \sqrt{\omega(\mathbf{u})}A(\mathbf{u})$  and  $k = 1, 2, \dots, M$ . *k* refers to the *k*<sup>th</sup> coordinate point pair and *M* refers to the number of pairs. k = v + (u - 1)h where *h* means the height of  $A(\mathbf{u})$ . The weight  $\omega(\mathbf{u})$  is denoted by,

$$\omega(\mathbf{u}) = Corr(\mathbf{u}) \cdot Dist(\mathbf{u}),\tag{9}$$

where  $Corr(\mathbf{u})$  refers to the correlation weight and  $Dist(\mathbf{u})$  refers to the distance one. They are defined as follows.

$$Corr(\mathbf{u}) = \frac{\mathbf{N}(B(\mathbf{u}))\mathbf{N}(B(\mathbf{0}))}{\||\mathbf{N}(B(\mathbf{u}))\|\||\mathbf{N}(B(\mathbf{0}))\||},\tag{10}$$

where  $B(\mathbf{u})$  denotes the square-block of data points with the side length *L* around the point **u** in the plane  $A(\mathbf{u})$ .  $B(\mathbf{0})$  corresponds to the point (0, 0).  $\mathbf{N}(B(\mathbf{u}))$  refers to the three-dimensional normal unit vector calculated by least-square

fitting of the data square-block  $B(\mathbf{u})$ . L = 10 is a typical and effective value in our experiments.

$$Dist(\mathbf{u}) = 1 - \sqrt{2[(u/w)^2 + (v/h)^2]},$$
(11)

where *w* and *h* refer to the width and height of  $A(\mathbf{u})$ . The range of point **u** is from (-w/2, -h/2) to (w/2, h/2). If the point lies at the center of the linear phase plane, e.g.,  $\mathbf{u} = (0, 0)$ , the distance weight gets the high value; if the point is one of the diagonal points in the plane, e.g.,  $\mathbf{u} = (w/2, h/2)$ , the distance weight gets the low value.

In our scheme, we assume that the data points in the center of the plane (corresponding to the low-frequency band) are more reliable than others. Our assumption is reasonable because outliers on the low-frequency band are much less than ones on the high-frequency band [8]. According to that, we calculate the correlation between each data block and the central block. The higher results lead to higher confidence-weight, and vice versa. In addition, the low-frequency points are considered more effective and their confidence-weights are higher.

# 4. Experiments and Discussion

Both the synthetic and real data are tested on the performance of the accuracy and robustness against noise. Let  $d_x(i)$  and  $d_y(i)$  denote the absolute registration errors along the x- and y- directions, i.e.,  $d_x(i) = |x_i - \hat{x}_i|$  and  $d_y(i) = |y_i - \hat{y}_i|$ , where  $(x_i, y_i)$  are the *i*<sup>th</sup> ground truth and  $(\hat{x}_i, \hat{y}_i)$ are their corresponding estimations. The MAE (Mean of the Absolute Errors) and the SAE (Standard deviation of the Absolute Errors) are used as the measurement, which are defined by,

$$MAE(z) = \frac{1}{n} \sum_{i=1}^{n} d_z(i),$$

$$SAE(z) = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (d_z(i) - MAE(z))^2},$$
 (12)

where *n* refers to the times of simulations and z = x, y. First we take experiments on different combinations of our modules. Then our scheme is compared with the schemes of Patrick [8] and Keller [9] implemented by us.

# 4.1 Modules Evaluation

In this section, modules of our framework are evaluated, including the denoising and confidence-weighted optimization modules. We take experiments on 4 images shown in Fig. 2. For each image, a test pair with the shifts of  $\Delta \mathbf{x} = (0.5, 0.5)$  is created. Then the White Gaussian Noise (WGN) with different variances is added to the test pairs and the variances are from 0 to 12, 000 with the interval of 1, 000. This procedure is repeated 100 times. Four combinations of our modules are compared, i.e., PC (Phase Correlation), PC+ED (Edge-preserving Denoising), PC+CO (Confidence-weighted Optimization) and PC+ED+CO (the proposed method). The MAE with the SAE along the x-and y- directions versus the noise variance are plotted in Fig. 3 (a) and (d) respectively.

We draw three conclusions from the above experiments. 1) As the EF and CO modules filter off outliers partly







**Fig.3** The MAE with the SAE (y-axis) versus the noise variance (x-axis). (a) and (d) refer to the evaluation of modules along the x- and y- directions. (b) and (e) refer to the evaluation of schemes for the synthetic data along the x- and y- directions. (c) and (f) refer to the evaluation of schemes for the real data along the x- and y- directions.

in the linear phase plane, the estimations are more accurate than PC. 2) The CO module proves to be more effective than the EF one. That's because giving low weights to the data points on the high-frequency band can alleviate the noise. 3) The proposed method achieves the most accurate and robust results among all.

# 4.2 Schemes Evaluation

Firstly, 10 standard images in the literature are employed as the synthetic data. First, in each direction 9 shifts are generated within the interval of [-1, 1], which include -1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75 and 1. Consequently, for each image in total 81 2-D shifts are formed to generate sub-pixel samples. Then the additive WGN with different variances is embedded on those samples and the procedure is repeated 100 times. The corresponding variances are from 0 to 40, 000 with the interval of 1, 000. When the variance is larger than 12, 000, the MAE of all the methods are larger than 1 which are unreliable. So Fig. 3 (b) and (e) only show the MAE with the SAE along the x- and y- directions versus the noise variances from 0 to 12, 000.

Next, real MRI (Magnetic Resonance Imaging) data is evaluated. The dataset is from Hoge [2], containing a series of MRI images, and a sample is shown in Fig. 2 (d). The real shifts between each pair of images are known. After establishing image pairs from given data, the additive WGN with different variances is embedded on images for 100 times. The corresponding variances are from 0 to 40, 000 with the interval of 1, 000. In this experiment, integer parts have been compensated firstly. When the variance is larger than 6, 000, the MAE of all the methods are larger than 1 which are unreliable. Thus only the MAE with the SAE along the x- and y- directions versus the noise variances from 0 to 6, 000 are shown in Fig. 3 (c) and (f).

Some conclusions can be drawn from the above results. 1) All methods report good results in the high SNR. The reason why the estimations of the real data along the x- direction are bad is analyzed in [2]. 2) The MAE of all methods is less than 0.01 under the condition that images registered are noisy-free, i.e. variance=0. Further more, we achieve the minimal MAE among all. 3) With the SNR decreasing, our method becomes more accurate and robust against noise than Patrick's and Keller's, especially in the experiments for the synthetic data.

#### 5. Conclusion

In summary, we propose a high-accuracy sub-pixel registration framework which is composed of the denoising and optimization modules. In the denoising module, we acquire a more original image signal by applying an edge-preserving filter. In the optimization module, we fit the confidenceweighted plane to achieve sub-pixel shifts. Experimental results on both synthetic and real data demonstrate that our framework is reasonable and achieves the most accurate and robust estimations among recent research works.

#### References

- B. Zitova and J. Flusser, "Image registration methods, a survey," Image Vis. Comput., vol.21, no.11, pp.977–1000, Nov. 2003.
- [2] W.S. Hoge, "Subspace identification extension to the phase correlation method," IEEE Trans. Med. Imaging, vol.22, no.2, pp.277–280, Feb. 2003.
- [3] H.S. Stone, M. Orchard, E.-C. Chang, and S. Martucci, "A fast direct Fourier-based algorithm for sub-pixel registration of images," IEEE Trans. Geosci. Remote Sens., vol.39, no.10, pp.2235–2243, Oct. 2001.
- [4] C.D. Kuglin and D.C. Hines, "The phase correlation image alignment method," Proc. Int. Conf. Cybernetics Society, 1975, pp.163– 165.
- [5] H. Foroosh, J.B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," IEEE Trans. Image Process., vol.11, no.3, pp.188–200, 2002.
- [6] V. Argyriou and T. Vlachos, "A study of sub-pixel motion estimation using phase correlation," Proc. BMVC, 2006, pp.387–396, 2006.
- [7] J. Ren, J. Jiang, and T. Vlachos, "High-accuracy sub-pixel motion estimation from noisy images in Fourier domian," IEEE Trans. Image Process, vol.19, no.3, pp.1379–1384, 2010.
- [8] P. Vandewalle, S. Ssstrunk, and M. Vetterli, "A frequency domain approach to registration of aliased images with application to superresolution," EURASIP J. Appl. Signal Process, vol.2006, pp.1–14, 2006.
- [9] Y. Keller, A. Averbuch, and O. Miller, "Robust phase correlation," Proc. CVPR, pp.740–743, 2004.
- [10] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," ICCV, 839.846, 1998.