Robust Phase Correlation based Motion Estimation and Its Applications

Hongshi Yan and Jian Guo Liu

Department of Earth Science and Engineering
Imperial College London
Prince Consort Road, London SW7 2AZ

{h.yan, j.g.liu}@imperial.ac.uk

Abstract

This paper presents a robust phase correlation technique and a compound phase correlation method for reliable sub-pixel image feature matching and motion estimation. A phase fringe filter and a highly robust estimator QMDPE are used to improve the fitting accuracy of the phase difference plane in Fourier domain. A compound phase correlation method is proposed to identify and decompose the multiple motion patterns in phase difference matrix at sub-pixel accuracy. This compound phase correlation method is the breakthrough in improving the robustness of phase correlation based feature matching around motion boundary or depth discontinuity. With the robust phase correlation and the compound phase correlation combined algorithm, the optical flow estimation and stereo matching achieved remarkable accuracy, especially around the areas with the motion boundary or depth discontinuity.

1 Introduction

Motion provides rich information for understanding the scene in an image sequence. An accurate estimation of motion in an image sequence is a crucial step in many computer vision and image processing applications such as image registration, multi-frame super resolution and stereo matching. In recent years, phase correlation feature matching method has been a popular choice in estimating the global or local translational motions between two similar images due to its remarkable accuracy and its robustness to uniform variations of illumination and signal noise in images [1].

The original phase correlation method [2] is known to identify integer pixel displacement. Several Fourier domain methods [3] [4] [8] and closely related spatial domain variations [1] have been proposed for estimating the translational shift with sub-pixel accuracy between image pairs. Stone *et al.* [3] investigated the effects of aliasing on the shift estimation and proposed a direct Fourier-based algorithm for sub-pixel image registration, in which, the translational parameter is directly estimated in Fourier domain through a least-squares fitting (LSF) to a 2D phase difference data set. Foroosh

BMVC 2008 doi:10.5244/C.22.104

et al. [1] claimed that Stone's approach is rather inaccurate since it often requires unwrapping the noisy 2D phase difference data and then fitting the unwrapped data. Alternatively, they extended the original phase correlation method [2] to sub-pixel accuracy through analytic expressions for the phase correlation of down-sampled images. Hoge [4] demonstrated that the translational shift between two images can be obtained by finding the rank-one approximation of the phase correlation matrix through the singular value decomposition (SVD) method. Then, the sub-pixel estimates of vertical and horizontal shifts can be derived independently from the left and right singular vectors. The accuracy of the translational shift estimation through the SVD based rank-one approximation of phase correlation matrix is much higher than Foroosh *et al.*'s method [1]. However, the computation complexity of the SVD operation of a large size matrix is very high.

Motivated by the strengths and limitations of these existing phase correlation methods for sub-pixel translational motion estimation and registration, this paper first addresses a robust phase correlation technique achieving reliable sub-pixel image feature matching and motion estimation at high speed. We propose to use a phase fringe filter [5] and Quick Maximum Density Power Estimator (QMDPE) robust techniques [6] in the direct Fourier-based phase correlation algorithm. We first apply the phase fringe filter to reduce the noise in the phase correlation difference matrix, and thus make the 2D unwrapping reliable. We then use the highly robust QMDPE technique to obtain the best fitting estimation of 2D unwrapped phase plane.

Phase correlation techniques are often applied locally for motion flow estimation and disparity estimation in stereo matching [7] [8]. However, its degraded performance around motion boundaries or depth discontinuity areas is well recognised [9] [10], which is also a challenge to most of the existing motion estimation methods. Noting the problem of phase correlation based motion estimation around motion boundary or depth discontinuity areas, this paper also investigates the characteristics of phase correlation of the image pair with complicated motion property. We proposed and designed a compound phase correlation (CPC) method to identify and decompose the multiple motion patterns in phase correlation matrix at sub-pixel accuracy, which is the breakthrough in improving the robustness of phase correlation based feature matching around motion boundary or depth discontinuity.

This paper is organised as follows. In the next section, we describe a direct Fourier based robust phase correlation technique for translational motion estimation. In Section 3, we address the compound phase correlation method for multiple motion decomposition and motion estimation with sub-pixel accuracy. Section 4 presents the combined phase correlation based algorithm for optical flow estimation and disparity estimation for stereo matching. Experimental results are then given in Section 5. Finally in Section 6 some concluding remarks are provided.

2 Robust Phase Correlation Technique

Phase correlation provides straight-forward estimation of rigid translational motion between two images, which is based on the well-known Fourier shift property: a shift in the spatial domain of two images results in a linear phase difference in the frequency domain of the Fourier Transforms (FT). Given two 2D functions $g(x,y)$ and $h(x,y)$ representing two images related by a simple translational shift a in horizontal and b in vertical directions, and the corresponding Fourier Transforms are denoted $G(u,v)$ and $H(u,v)$. Thus,

$$H(u,v) = G(u,v)\exp\{-i(au+bv)\} \tag{1}$$

The phase correlation is defined as the normalised cross power spectrum between G and H, which is a matrix:

$$Q(u,v) = \frac{G(u,v)H(u,v)^*}{\left|G(u,v)H(u,v)^*\right|} = \exp\{-i(au+bv)\} \tag{2}$$

If $G(u,v)$ and $H(u,v)$ are continuous functions, then the inversed Fourier Transform (IFT) of $Q(u,v)$ is a Delta function. The Delta function peak identifies the integer magnitude of the shift between the pair of images [2]. To achieve the translation estimation at sub-pixel accuracy, a common approach is to oversample images $g(x,y)$ and $h(x,y)$ to sub-pixel level before the FT of phase correlation operations. This however will increase the computing load dramatically. Many researchers looked for a direct solution in frequency domain based on the phase correlation matrix defined in (2). As the magnitude of $Q(u,v)$ is normalised to 1, the only variable in (2) is the phase difference defined by $au+bv$, where a and b are the horizontal and vertical magnitudes of the image shift between $g(x,y)$ and $h(x,y)$. If we can solve a and b accurately based on the phase correlation matrix $Q(u,v)$, then the non-integer translation estimation at sub-pixel accuracy can be achieved without applying IFT. Such direct frequency domain approaches [3] [4] has been proved more accurate and faster than that based on the Delta function method.

The phase difference angle $c= au+bv$ in (2) is simply a planar surface through the origin in u-v coordinates defined by coefficients a and b. Thus a complicated problem of complex numbers in frequency domain becomes a simple issue of finding the best 2D fitting of the phase difference angle data in $Q(u,v)$ to a plane of phase difference in the coordinates of u and v. The phase shift angle c is 2π wrapped in the direction defined by a and b. Any a 2D fitting for c is not possible without a 2D unwrapping. However, 2D unwrapping on the phase angle data in the $Q(u,v)$ is often unreliable and results in failure of finding a and b correctly [1] [4]. This is largely because of the noisier data of $Q(u,v)$. To improve the 2D fitting method, we recognized that the key issues are: to reduce the data noise before unwrapping and to refine the fitting technique.

As the phase angle data in the $Q(u,v)$ is 2π wrapped, ordinary smoothing filters cannot be applied directly to reduce the noise of such discontinuous periodical data. We implemented a phase fringe filtering technique [5] into the 2D fitting method as below:

1. Denote $\theta(u,v)$ as the phase angle at position u,v in the phase correlation matrix $Q(u,v)$.
2. The $\sin\theta$ and $\cos\theta$ are continuous functions of $\theta(u,v)$, a smoothing filter can therefore be applied to these functions.
3. Derive the filtered phase angle $\overline{\theta(u,v)}$ from smoothing filtered $\sin\theta$ and $\cos\theta$:

   $\tan\bar{\theta} = \dfrac{\overline{\sin\theta}}{\overline{\cos\theta}}$.

The window size of the smoothing filter used must be small in comparison to the half wavelength of $\sin\theta$ and $\cos\theta$. For reducing the aliasing error and edge effects in the direct Fourier based method, high frequency components of the phase correlation matrix should be masked out, and only the lower frequency part is kept for the 2D fitting operation [3] [4].

In stead of using LSF, a robust fitting technique QMDPE [6] is finally applied to find the best fitting estimates of the unwrapped phase angle data, which often is contaminated by the incorrectly unwrapped data and contain multi-structure mode. It is noticed that the robust estimator QMDPE can tolerate more than 80% of outliers, and it has been proved more robust than another common robust estimator RANSAC [6]. The QMDPE method repeatedly selects a random set of three points within the unwrapped phase angle data and gets the plane transformation model induced by them. The plane model with the largest density power value can be obtained through a mean shift procedure, and such set of unwrapped phase angle data is chosen to represent the statistical inliers. A least-squares solution finally applies to these inliers to form the final estimates of the phase angle plane. The benefit of using the QMDPE robust estimator is that the best estimate of the translational shifts a and b in (2) can be obtained from the noisy phase difference data set.

An example in Figure 1 shows the effectiveness of the robust phase correlation based translational motion estimation method for sub-pixel image registration across different spectral bands. A image pair of two different spectral bands, bands 1 (blue) and 5 (short wave infrared), extracted from a 30 m resolution Landsat-7 ETM+ scene is shown in Figure 1(a) and 1(b) respectively. The correlation between the two bands is 0.69. One of the images is shifted horizontally by 13.33 pixels to the left and vertically by 10.00 pixels up in relation to the other. As shown in Figure 1(c), the phase correlation matrix data become quite noisy because of the low correlation between the two images. Both of the Least-Square Fitting (LSF) and QMDPE algorithms failed in the first attempt without filtering the noise phase angle data. Then the phase fringe filter with filter size 15×15 pixels was applied, which has improved the phase correlation data significantly as illustrated in Figure 1 (d) and 1(e), and ensures a successful 2D unwrapping result shown in Figure 1(f). Only after the filtering, the LSF and robust QMDPE algorithms succeeded in measuring the image frame shift at sub-pixel accuracy. As comparison, the phase fringe filter with small size 5×5 pixels was also applied and the filtered result is shown in Figure 1(g). The filtered result is not as good as that in Figure 1(d), and some errors exist in the corresponding unwrapped phase angle data shown in Figure 1(h). However, the QMDPE robust fitting method still obtained very good estimates. LSF cannot find the good fitting estimates in this case. The experimental results shown in Table1 and Table 2 indicate that the QMDPE 2D fitting algorithm achieved the best overall accuracy with different fringe filter size.

True Shifts	LSF	QMDPE
x: 13.3333	13.2016	13.2056
y: -10.00	-9.9324	-9.9404

Table 1: Translational shift estimates with fringe filter size 15×15.

True Shifts	LSF	QMDPE
x: 13.3333	13.2466	13.2420
y: -10.00	- 9.3308	- 9.9346

Table 2: Translational shift estimates with fringe filter size 5×5.

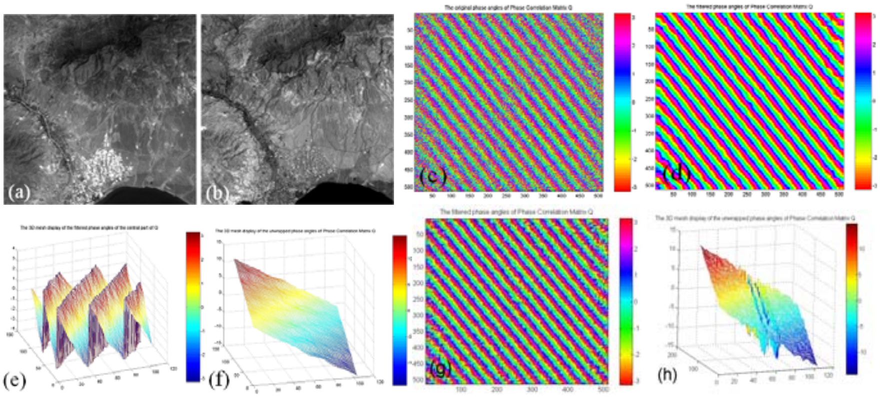

Figure 1: Image (a) and (b) are Landsat-7 ETM+ image bands 1 and 5 with inter band correlation of 0.69. Image (b) is artificially shifted to the left by 13.3333 pixels and up by 10 pixels. (c) The phase difference matrix. (d) The filtered phase difference matrix (filter size=15). (e) 3D view of the central part of the filtered phase difference matrix (filter size=15). (f) 3D view of the unwrapped phase difference matrix (filter size=15). (g) The filtered phase difference matrix (filter size=5). (h) 3D view of the central part of the filtered unwrapped phase difference matrix (filter size=5).

3 Compound Phase Correlation based Multiple Motion Decomposition

The phase correlation difference matrix around an area with multiple motions tends to be rather messy with overlaid and interfered multiple fringe patterns, which cannot be unwrapped properly. Our investigation indicates that these overlaid fringe patterns with different orientations and frequencies correspond to the each motion mode within this area respectively. Identification and decomposition of these patterns thus enable accurate local motion estimation around motion boundaries or depth discontinuities.

It is a difficult task to identify and decompose the multiple motion patterns directly in the phase correlation matrix. However, if we apply the inverse Fourier transformation to the phase correlation matrix, the multiple motions can be clearly separated as different Delta function impulses. The corresponding motion can then be easily estimated based on the locations of these Delta function impulses, but only at a pixel level accuracy. This is so called Delta function based phase correlation method [2]. Foroosh *et al.* [1] proved that a peak of the Delta function determines the integer part of the corresponding motion while the close neighborhood of the peak determines the sub-pixel part of the motion. In order to accurately and robustly estimate the corresponding motions, we designed a compound phase correlation (CPC) method which benefits both advantages of easy decomposition of multiple motions from Delta function method in spatial domain and sub-pixel accuracy of the direct phase correlation technique for individual motion estimation in Fourier domain.

For each decomposed motion, we should mask out all other parts of the 2D Delta function impulse and only keep the immediate neighbourhoods of the dominant peaks, and then apply Fourier transformation to each peak and its immediate neighbour to obtain a sub-matrix of phase correlation that corresponds only to one dominant motion. Finally, the simple 2D LSF fitting algorithm is applied to each of the phase correlation sub-matrices to obtain of the multiple motion estimations with sub-pixel accuracy. We can simply repeat the above steps to achieve accurate estimate of each of the multiple motions within a large scanning window. It should be noted that the general CPC method requires linking each Delta function peak to its corresponding motion, which is not an easy task if the scanning window contains complex motion modes. For the reason of easy operation and algorithm efficiency, we simply presume that the motion at the centre of each window can be derived from the first dominant motion mode. In practice, this simple approach achieves good performance in most cases except around some corners of multiple motion boundaries, which is also a big challenge to most exiting motion estimation techniques.

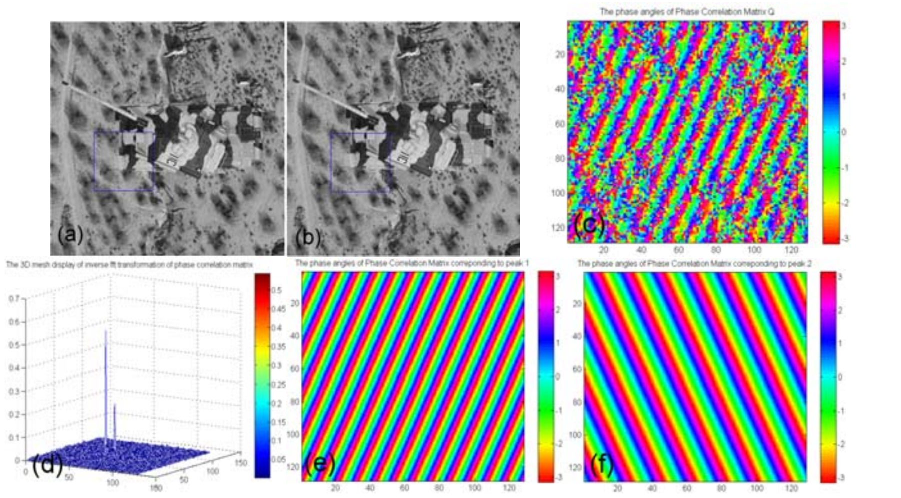

Figure 2: Motion decomposition through the proposed CPC method. An image pair (a) and (b) show a moving tank in a background in motion. (c) The phase correlation matrix between the two blue rectangle areas in the image pair. (d) Two distinctive peaks of the Delta function that correspond to the tank motion and the background motion respectively. (e) The decomposed phase correlation matrix of the moving tank. (f) The decomposed phase correlation matrix of the moving background.

Figure 2 shows an example of motion decomposition around motion boundary through the proposed CPC method. We inserted a picture of a tank model into different positions in a sand desert image pair, which is generated by artificially shift one image by 7 pixels (to left) in horizontal and 3 pixels (down) in vertical directions. The tank

	Tank	Background
True shifts	(-12, -5)	(-7, 3)
Estimates (128×128)	(-11.9723, 4.9621)	(-6.9060, 2.9754)
True shifts	(-6, -2.5)	(-3.5, 1.5)
Estimates (64×64)	(-5.984, -2.572)	(3.3589, 1.681)

Table 3: Results of multiple motion estimation through CPC technique

position change between Figure 2(a) and Figure 2(b) is 12 pixels (to left) horizontally and 5 pixels (up) vertically. We chose a window area around motion boundary (blue rectangle) in each of the image pair for the motion decomposition test. The size of the blue rectangle is 128×128. The phase correlation matrix between the two blue rectangle areas in the image pair is shown in Figure 2(c), while Figure 2(d) illustrates that the inverse Fourier transformation of the phase correlation matrix resulted in two distinctive Delta function peaks corresponding to the tank motion and the background motion respectively. Figure 2(e) and 2(f) show the decomposed phase correlation matrices of the moving tank and the background. The motions of tank and the background were accurately estimated from the clean and sharp fringe patterns of the corresponding phase difference matrix using the simple LSF phase correlation technique. In addition, we also down-sampled the image pair by two for further test around the same motion boundary area but with window size 64×64. The motion estimates of tank and background are shown in Table 3. This example indicates that the CPC method is able to achieve multiple motion decomposition and estimation at sub-pixel accuracy.

4 Combined Phase Correlation based Local Motion Estimation

There are some limitations to apply the proposed robust 2D fitting based phase correlation method locally for optical flow estimation. Here, the scanning window size is crucial for the quality of the motion estimation. If the window size is too small, then the number of data points will be insufficient to achieve accurate measurement of the extracted feature shift, while if it is too large it may include multiple motions, especially around depth discontinuity areas for stereo matching.

The proposed novel CPC method has much better performance for optical flow estimation around motion boundaries or depth discontinuity than the original robust 2D fitting technique, but it is generally less efficient and slightly less accurate in normal areas without motion boundaries or depth discontinuity. So we designed a combined motion (disparity) estimation algorithm that employs the corresponding advantages of the both techniques. The key issue of this newly designed combined motion estimation algorithm is to locate the motion boundaries where the phase correlation data are messy, and apply the CPC for motion (disparity) estimation only in these selected areas while keeping the original QMDPE based robust technique as the default phase correlation engine for most parts of the image. This simple and effective automatic processing procedure comprises two steps:

1. Carry out the raster scan of the image pair using a moving window to estimate the motion of each small window with the robust QMDPE phase correlation algorithm.
2. In each window, the robust fitting technique QMDPE is applied to find the best fitting estimates of the unwrapped phase angle data. However, the estimation would be poor around motion boundaries where the phase angle data are contaminated by the incorrect unwrapping and multi-structure mode because of poor correlation or high noise level. If the ratio of the outliers to the inliers of the best fitting estimation of plane model exceeds a certain threshold, the corresponding motion estimate is supposed to be poor, and then the CPC technique is used to determine the dominant motion within the phase correlation window.

In the step two above, we proposed to use a ratio of outliers to inliers derived from the robust QMDPE 2D fitting in frequency domain for quality assessment on the phase correlation based motion estimation. Beside this, the linear regression correlation coefficient between the actual phase difference data and the robust estimates of the phase difference plane can also be employed to assess the quality of the phase correlation based motion estimation.

5 Experimental Results

Figure 3(a) and 3(b) show a pair of tank images generated in a similar way as described in Figure 2, which simulate a vertical view sequence of a moving military target in a sand desert background (no motion). The tank in the image on the right is shifted 2.5 pixels to the left and 3 pixels upward. In order to test the robustness of the proposed method to illumination variations, both of the brightness and contrast of the image pair have been changed. The image pair was scanned with 32×32 window to compute the optical flow field. The estimated optical flow field through the robust 2D fitting phase correlation technique without motion decomposition is shown in Figure 3(c) and its corresponding magnitude image is shown in Figure 3(f). Figure 3(d) shows that the low quality optical flow data in Fig 3(c) have been masked off with the robust inliers scale estimation method, and can be refilled with the proposed PCP technique. The estimated optical flow field using the robust 2D fitting and CPC combined algorithm is shown in Figure 3(e) and its corresponding magnitude image is shown in Figure 3(g), which indicate that the proposed robust 2D fitting and CPC combined method has successfully computed the optical flow field with sharp motion edge from an image pair with illumination variation, which is a big challenge to most existing techniques.

One of important applications of CPC combined optical flow estimation algorithm is to improve the accuracy of disparity measurement across steep slopes in very narrow baseline stereo image pair for stereo matching.

One image of a stereo pair of images from CMU Castle sequence is shown in Figure 4(a). The disparity map generated using the CPC based combined approach is shown in Figure 4(b). For comparison, the disparity map from our robust phase correlation technique without motion decomposition is shown in Figure 4(c). The test results are self evident; the new method has obviously improved performance, especially in disparity estimation around the sharp edges of buildings and walls. Figure 4(d) is a 3D perspective view of Castle image reconstructed from the estimated disparity map. It demonstrates that fine details and depth discontinuity can be quite effectively recovered.

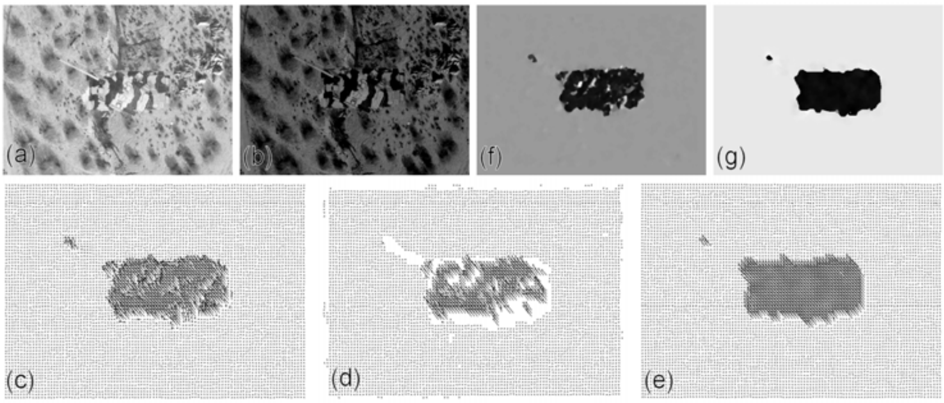

Figure 3: (a) and (b) An image pair with a moving tank in a still desert background. (c) The computed optical flow field from the robust 2D fitting technique. (d) The low quality optical flow data in (c) have been masked off with the robust inliers scale estimation method. (e) The estimated optical flow field through our robust 2D fitting motion estimation and CPC combined algorithm. (f) The magnitude image of the optical flow field (c). (g) The magnitude image of the flow field (e).

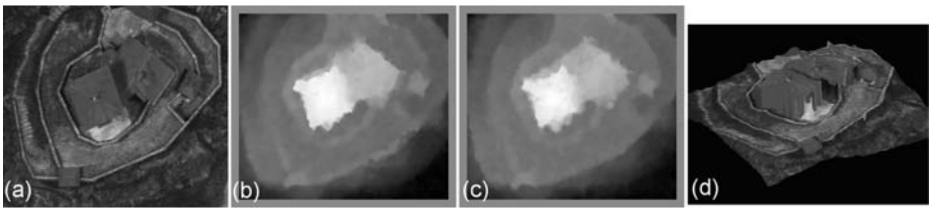

Figure 4: Experiment results of disparity estimation from a stereo image pair of CMU. (a) One image of the stereo image pair from CMU Castle sequence. (b) The disparity measurement image produced by the CPC based combined approach. (c) The disparity measurement image produced by the simple robust phase correlation technique. (d) The 3D perspective view of Castle image based on the CPC derived disparity measurement image in (b).

6 Conclusion

The analyses and experimental results presented in this paper have demonstrated that the proposed phase correlation based motion estimation scheme, which combines the robust phase correlation technique and CPC method, are able to achieve remarkable accuracy in most synthetic and real images from different spectral bands. The strengths of the proposed technique are its algorithm simplicity, its robustness to illumination change and its good performance around motion boundaries or depth discontinuities.

Acknowledgements

The work reported in this paper was funded by the Systems Engineering for Autonomous Systems (SEAS) Defence Technology Centre established by the UK Ministry of Defence.

References

[1] H. Foroosh, J. B. Zerubia, and M. Berthod, “Extension of phase correlation to subpixel registration,” IEEE Trans. Image Processing, vol. 11, no. 3, pp. 188–200, Mar. 2002.

[2] C. D. Kuglin and D. C. Hines, “The phase correlation image alignment method,” *Proceeding of IEEE International Conference on Cybernetics and Society*, pp. 163–165, New York, NY, USA, September 1975.

[3] H. S. Stone, M. T. Orchard, E.-C. Chang, and S. A. Martucci, “A fast direct Fourier-based algorithm for subpixel registration of image,” *IEEE Transactions on Geoscience and Remote Sensing,* vol. 39, no. 10, pp. 2235–2243, 2001.

[4] W. S. Hoge, “Subspace identification extension to the phase correlation method,” *IEEE Trans. Medical Imaging*, vol. 22, no. 2, pp. 277–280, 2003.

[5] Wang Feng, Prinet Veronique and Ma Songde, “A vector filtering technique for SAR interferometric phase image,” http://kepler.ia.ac.cn/publications/2001/Wangfeng

[6] H. Wang and D. Suter, “A very robust estimator for modelling fitting and range image segmentation,” *International Journal and Computer Vision*, vol.59, no. 2, pp. 139-166, 2004.

[7] D. J. Fleet, “Disparity from local weighted phase-correlation,” *Proc. IEEE International Conference on Systems, Man and Cybernetics*, pp. 48-56, October 1994.

[8] M. Balci and H. Foroosh, “Inferring motion from the rank constraint of the phase matrix,” *IEEE ICASSP 2005 Proc.,* Vol. II, pp. 925-928, 2005.

[9] S. S. Beauchemin, K. Daniilidis and R. Bajcsy, “Computing multiple image motions,” *Vision Interface*, pp. 272-279, 1998.

[10] Tweed and A. Calway, “Estimating dense motion fields using a local two component model,” *Proc. IEE Conference on Image Processing and its Applications*, pp. 696-700, 1999.