A Novel Subpixel Phase Correlation Method Using Singular Value Decomposition and Unified Random Sample Consensus

Xiaohua Tong, Zhen Ye, Yusheng Xu, Shijie Liu, Lingyun Li, Huan Xie, and Tianpeng Li

***Abstract*—Subpixel translation estimation using phase correlation is a fundamental task for numerous applications in the remote sensing community. The major drawback of the existing subpixel phase correlation methods lies in their sensitivity to corruption, including aliasing and noise, as well as the poor performance in the case of practical remote sensing data. This paper presents a novel subpixel phase correlation method using singular value decomposition (SVD) and the unified random sample consensus (RANSAC) algorithm. In the proposed method, SVD theoretically converts the translation estimation problem to one dimensions for simplicity and efficiency, and the unified RANSAC algorithm acts as a robust estimator for the line fitting, in this case for the high accuracy, stability, and robustness. The proposed method integrates the advantages of Hoge's method and the RANSAC algorithm and avoids the corresponding shortfalls of the original phase correlation method based only on SVD. A pixel-to-pixel dense matching scheme on the basis of the proposed method is also developed for practical image registration. Experiments with both simulated and real data were carried out to test the proposed method. In the simulated case, the comparative results estimated from the generated synthetic image pairs indicate that the proposed method outperforms the other existing methods in the presence of both aliasing and noise, in both accuracy and robustness. Moreover, the pixel locking effect that commonly occurs in subpixel matching was also investigated. The degree of pixel locking effect was found to be significantly weakened by the proposed method, as compared with the original Hoge's method. In the real data case, experiments using different bands of ZY-3 multispectral sensor-corrected images demonstrate the promising performance and feasibility of the proposed method, which is able to identify seams of the image stitching between sub-charge-coupled device units.**

***Index Terms*—Aliasing and noise, singular value decomposition (SVD), subpixel phase correlation, unified random sample consensus.**

Manuscript received June 2, 2014; revised October 31, 2014 and December 6, 2014; accepted January 11, 2015. This work was supported in part by the National Natural Science Foundation of China under Project 41325005, Project 41401531, and Project 41171352; by the National High-tech Research and Development Program under Project 2012AA120905; by the National Key Basic Research Program of China (973 Program) under Project 2012CB957701; and by the China Special Funds for Meteorological and Surveying, Mapping and Geoinformation Research in the Public Interest under Project GYHY201306055 and HY14122136.

The authors are with the College of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, China (e-mail: xhtong@tongji.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2015.2391999

I. Introduction

PHASE correlation is a Fourier-based matching technique, which is considered to be more accurate and effective than the commonly used correlation methods such as normalized cross correlation [1], [2]. Phase correlation has been successfully adopted in a large number of applications, such as super-resolution reconstruction [3], medical image registration [4], motion tracking [5], digital image stabilization [6], mosaicking [7], fingerprint matching [8], 3-D construction [9], and video analysis and processing [10], [11]. In addition, phase correlation has also drawn more and more attention in the remote sensing community, in applications such as pixel-to-pixel coregistration [12], narrow baseline digital elevation model (DEM) generation [13], [14], in-flight calibration [15], [16], and surface dynamics measurement, including coseismic deformation measurement [17]–[19], glacier displacement survey [1], [20], and sand dune migration investigation [21], [22]. This paper concentrates on introducing subpixel image registration for translation estimation, which is fundamental to many applications in the fields of remote sensing and image processing, such as earth surface dynamics monitoring [23], [24], attitude jitter analysis [25], [26], image super-resolution [27], and optical flow [28].

All the current phase correlation methods are based on the well-known Fourier shift property, which states that a shift of two relevant images in the spatial domain is transformed into the Fourier domain as linear phase differences [29]. Two image functions that are related by shifts x_0 and y_0 in the line and column directions are denoted as $f(x, y)$ and $g(x, y)$, and expressed as

$$g(x, y) = f(x - x_0, y - y_0). \tag{1}$$

According to the Fourier shift property by means of the Fourier transform (FT)

$$G(u, v) = F(u, v) \exp\{-i(ux_0 + vy_0)\} \tag{2}$$

where $F(u, v)$ and $G(u, v)$ are the corresponding FT of $f(x, y)$ and $g(x, y)$. The normalized cross-power spectrum matrix is given by

$$Q(u, v) = \frac{F(u, v)G(u, v)^*}{|F(u, v)G(u, v)^*|} = \exp\{-i(ux_0 + vy_0)\} \tag{3}$$

where * denotes the complex conjugate, and Q is the normalized complex matrix that only relates to the phase shift

components of the two input images. It is therefore insensitive to illumination differences and image content [13]. In the case of integer pixel shifts, the inverse Fourier transform (IFT) of $Q(u, v)$ is a Dirac delta function centered on (x_0, y_0), if $F(u, v)$ and $G(u, v)$ are supposed to be continuous functions (replaced by a unit impulse in the discretized case) [2]. Therefore, integer pixel shifts can be estimated from the peak coordinates of the IFT of the normalized cross-power spectrum matrix [29].

The existing researches indicate that the subpixel phase correlation methods can be generally classified into two categories [2], [17]. The first category estimates the translational parameters by precisely determining the main peak location of the IFT of the normalized cross-power spectrum. The second category calculates the relative displacement by estimating the linear phase difference directly in the Fourier domain.

In the former category, the most commonly used approach to estimate the subpixel offset is based on interpolation. In [30]–[32], interpolation methods with a 1-D parabolic function, Gaussian function, sinc function, and modified esinc function in the x- and y-directions, respectively, were adopted with three points, including the main integer peak and its two neighboring side peaks. However, these subpixel methods based on simple interpolation are sensitive to noise and other sources of error [33], and their accuracy is highly dependent on the data quality and interpolation algorithms [2]. In [2], Foroosh *et al.* assumed that images with subpixel shifts were originally displaced by integer values on a high-resolution grid, which was followed by downsampling. According to [2], in the case of a subpixel shift, the signal power is not concentrated in a single peak but rather in several coherent peaks, with the most eminent ones largely adjacent to each other, which implies that phase correlation leads to a downsampled 2-D Dirichlet kernel. Compared with the methods in [31] and [32], the method of Foroosh deduces a closed-form solution for the subpixel estimation by applying the sinc function to approximate the Dirichlet function, and is thus more analytically demonstrable. However, it has been claimed that Foroosh's method insufficiently takes into account the interference term during the analytical derivation. Using only one-sided information results in the method being subjected to the negative effect of noise [33], [34]. In [35], a so-called peak evaluation formula (PEF) derived from the sinc function fitting in one dimensions was introduced, through which the subpixel displacement can be achieved from multiple tri-tuples consisting of the main peak and its corresponding surrounding points, using least squares estimation without an iterative process. In [36], a matrix multiplication implementation of the 2-D discrete Fourier transform (DFT) was used to refine the conventional fast FT upsampling approach, and an initial estimation was obtained by the conventional method with an upsampling factor of 2. The final subpixel estimation by a defined upsampling factor was then computed in a 1.5×1.5 pixel neighborhood (in units of original pixels) around the initial estimation, using matrix-multiply DFT. Note that the method in [36] is free of an explicit measure to deal with aliasing and noise.

In the latter category, according to the fact that the phase shift angle (phase angle of $Q(u, v)$) is a linear function of the shift parameters, and represents a 2-D plane defined by shift parameters x_0 and y_0 through the origin of the u–v coordinates, Stone *et al.* [37] employed least squares adjustment to fit the 2-D plane with the removal of high-frequency and small-magnitude spectral components, which may be affected by aliasing and noise [3]. However, the phase shift angle is 2π wrapped in two dimensions, which means that only shifts of less than half a pixel can be measured. For this reason, two separate and consecutive 1-D unwrappings in the u-and v-directions were performed in [34]. A robust plane fitting approach, the quick maximum density power estimator, was applied in [13] and [38] to make the estimation more reliable. In [4], Hoge proposed a straightforward approach using SVD to find the dominant rank-one approximation of the normalized cross-power spectrum matrix before 1-D phase unwrapping. Keller [39] proposed a robust extension to [4] with a so-called "projection" masking operator, under the assumption that the noise is additive white Gaussian noise. However, this method achieves a poor improvement in practical applications. Nonlinear optimization is another commonly used direct approach for solving subpixel displacement. Leprince *et al.* [17] provided an effective method based on nonlinear optimization, and they minimized the Frobenius norm of the difference between the measured cross-spectrum and the theoretical one weighted by adaptive frequency masking, using a two-point step size gradient algorithm.

In this paper, a novel subpixel phase correlation method for translation estimation directly in the Fourier domain is proposed, with the aim being to improve the stability and robustness of the subpixel phase correlation method. The proposed method originates from Hoge's method [4], which is based on SVD, and is extended to be a more reliable and effective method by means of the RANSAC algorithm. The unified random sample consensus (RANSAC) algorithm, combining various modifications in different processing stages, is the core to our method, and provides the stability and robustness. Consequently, the proposed method integrates the advantages of Hoge's method and the RANSAC algorithm and avoids the corresponding shortfalls of Hoge's method. Subpixel image registration for translation estimation based on phase correlation was performed using simulated and real data to demonstrate the superiority and reliability of the proposed method. Quantitative comparisons between our method and the other prominent subpixel phase correlation methods using two sets of synthetic image pairs with individually increasing amounts of aliasing and noise were carried out. We also developed a pixel-to-pixel dense matching scheme based on the proposed method for use in practical applications, and we investigated the performance of our method using different bands of ZY-3 multispectral sensor-corrected (SC) images.

II. Proposed Phase Correlation Method

To obtain a subpixel accuracy, phase correlation methods that directly estimate the linear phase difference are more accurate and effective than identifying the location of the dominant peak, in a theoretical sense. As mentioned in [34], aliasing and noise artifacts, which are mainly localized at a high frequency in the Fourier domain, become dispersed in the spatial domain. It

was found that avoiding the IFT process contributes to dealing with aliasing and noise. Consequently, we favor translation estimation directly in the Fourier domain, and we propose a robust extension to Hoge's method [4] based on a unified RANSAC algorithm. This method allows us to acquire precise and reliable displacements at both the multipixel and subpixel scales, even in degraded conditions.

A. Original Hoge's Method

If the interference term is ignored, the normalized cross-power spectrum matrix Q is theoretically a rank-one matrix, which is shown in (3). $Q(u, v)$ is separable, which leads to the following equation:

$$\begin{aligned} Q(u,v) &= \exp\{-i(ux_0 + vy_0)\} \\ &= \exp\{-iux_0\}\exp\{-ivy_0\} = q_{x_0}(u)q_{y_0}(v). \end{aligned} \quad (4)$$

The problem of determining the subpixel shift parameters between two images is simplified by finding the best rank-one approximation of the normalized cross-power spectrum matrix Q [4]. According to the Eckart–Young–Mirsky theorem, the best low-rank approximation to a matrix is achieved by the SVD method [17]. Therefore, Hoge proposed to perform 1-D unwrapping on the phase angle of the left and right domain singular vectors independently after SVD process. The unwrapped phase angle of the domain singular vectors theoretically represents a straight line, and the subpixel shift parameters in the x- and y-directions can be estimated from the slope of the fitted line, respectively, using least squares estimation.

The idea of applying SVD to find the dominant rank-one subspace of the normalized cross-power spectrum matrix effectively avoids the ill-posed problem of 2-D phase unwrapping, which theoretically makes Hoge's method less sensitive to noise. Determining the dominant singular vectors implicitly acts as an eigenfiltering processing, which enhances the signal-to-noise ratio of the phase angle data and makes the phase unwrapping more reliable. Moreover, Hoge's method can be complementary with other subpixel phase correlation methods, and it possesses the ability to obtain integral and fractional shifts at the same time [4]. As a result of the aforementioned superiorities, Hoge's method, which has been reported to reach up to a 1/20th pixel accuracy, is widely used in many applications [4], [13], [39], [40]. It has also been reported that its performance on image pairs with large magnitude shifts or rotation, or low correlation, is better than the other common phase correlation methods [13].

However, although Hoge's method adopts an additional frequency masking operation [37], and good results have been reported in the original paper, its results become inaccurate in the presence of aliasing and noise, and its ability to handle these corruptions is thus questioned [17], [19], [33], [39], [41]. Furthermore, it is worth noting that the SVD computation is likely biased with unconsidered corrupted phase data due to noise and aliasing [17], making the method unreliable and unstable in the case of a small image size [13].

B. Unified RANSAC

A simple yet powerful robust estimator that does not require prior knowledge of outliers, namely, RANSAC, was proposed by Fischler and Bolles [42], and it greatly outperforms the standard least squares estimation in the presence of outliers. In the RANSAC algorithm, a subset of data with the minimum required size is randomly sampled. The model parameters are then estimated from this minimum size data and evaluated by checking whether the other data are inliers that support the model within a predefined threshold. The inliers of the best model so far constitute a consensus set. These procedures are repeated until the probability of finding a better model falls under the specified probability η_0 (typically set to 0.05). The iterative number k is defined according to the specified probability as $k = \log \eta_0 / \log(1 - \varepsilon^m)$, where ε denotes the fraction of inliers adaptively updated based on the size of the maximum consensus set found, and m is the minimum required size [43]. The final best model is chosen as the model with the maximum consensus set.

RANSAC is performed in a randomized, hypothesize-and-verify manner, which can yield good-quality estimates even when more than 50% of sample points are considered as gross errors [44], [45]. Although more than 30 years have passed, RANSAC-style algorithms are still one of the most widely used robust estimation methods, and they have been successfully applied in the computer vision and remote sensing community, in applications such as remote sensing image registration [46], [47], motion segmentation and estimation [48], [49], and epipolar geometry estimation [45], [50]. Moreover, a rich variety of refinements and modifications have also been proposed [45], [51]. In [45], Raguram *et al.* systematically reviewed recent research on the RANSAC-based robust estimator, and they refined each stage of the standard RANSAC algorithm by incorporating a number of practical and computational improvements and optimizations. RANSAC can be regarded as an optimization problem of minimizing a certain cost function. The original cost function of RANSAC is of the top-hat type, which is 0 for inliers and constant otherwise [51]. However, the top-hat cost function makes the RANSAC result unreliable when the predefined threshold for considering inliers is set unfavorably high. An enhanced cost function, which reflects how well the inliers fit the data, was therefore proposed in [52], according to the idea of a redescending M-estimator. Owing to the fact that drawing an all-inlier minimal random sample is insufficient to ensure a proper model hypothesis, a random sampling strategy with constraints and guidance can be adopted to accelerate and strengthen the RANSAC algorithm. In [53], a novel sampling strategy that aimed to draw all-inlier minimal samples with large spans was proposed, and a theoretical reasoning was presented. It was found in [54] that the actual iterative number was significantly higher than the theoretically predicted value, which was caused by the incorrect assumption that a model computed from an uncontaminated minimal sample is consistent with all the real noisy inliers. The authors in [54] thus adopted an local optimization processing once a better model is achieved during iteration. Four types of optimization strategies were presented, with the best model

so far as a starting point. The inner RANSAC with iteration method appears to be the optimal, wherein an inner RANSAC procedure based on a nonminimal random sample from inliers of the best model so far is iteratively performed. Iterative least squares with a decreasing threshold is further applied in each inner RANSAC procedure. The inner RANSAC procedure is carried out for ten iterations, and the refined model with the maximum consensus set is chosen.

In this paper, we integrate the effective ideas of [52]–[54] as a unified framework to extend the simple hypothesize-and-verify structure of the standard RANSAC, as recommended in [45]. The truncated quadratic cost function in the M-estimator SAC (MSAC) [52] replaces the top-hat type of RANSAC. In the minimal subset sampling stage, a prior sampling constraint is incorporated to prevent the minimal subset from containing data points that are too close, according to the principle in [53] that a minimal subset with a large spatial extent is desirable. A model refinement step, namely, inner RANSAC with an iteration method in locally optimized RANSAC [54], is inserted into the algorithm whenever a new best model so far is found. The use of the cost function in MSAC enhances the tolerance to the choice of the inlier threshold and makes it possible to select a suitable threshold for a diverse set of data [52]. A sampling strategy with a prior sampling constraint reduces the necessary number of iterations [53]. The additional local optimization step strengthens the precision, robustness, and stability of the algorithm [54]. It was demonstrated from the subsequent experiments and research in [45] that the unified RANSAC possesses the merits of accuracy, computational efficiency, stability of results, and insensitivity to parameter setting.

C. Proposed SVD-RANSAC Method

With regard to the drawbacks of Hoge's method aforementioned, an integrated and compact subpixel phase correlation method is introduced, which is referred to as the SVD-RANSAC method. In this context, we adopt the unified RANSAC algorithm to replace the least squares estimation method for the line fitting, as the RANSAC algorithm, particularly unified RANSAC, has many remarkably compelling advantages in robust estimation. Artifacts resulting from error sources, including aliasing and noise, are effectively excluded by the unified RANSAC algorithm. Therefore, SVD-RANSAC integrates the advantages of Hoge's method and the unified RANSAC algorithm and avoids the shortfalls of Hoge's method, and it can achieve subpixel translational estimates with a high reliability and strong robustness.

The overall workflow is depicted in Fig. 1 and described in the following. 1) The periodicity assumption of using DFT for finite-size discrete signals creates sharp discontinuities, inducing edge effects that produce spectral leakage in the frequency domain [2], [4]. The proposed method is able to handle slight edge effects, as discontinuities at the image border could be excluded by the unified RANSAC algorithm during line fitting. 2-D windowing functions weighting both input images are typically used to reduce edge effects [37]. In this case, we apply a raised-cosine window, which makes a reasonable tradeoff between reducing the artifacts and loss of information,

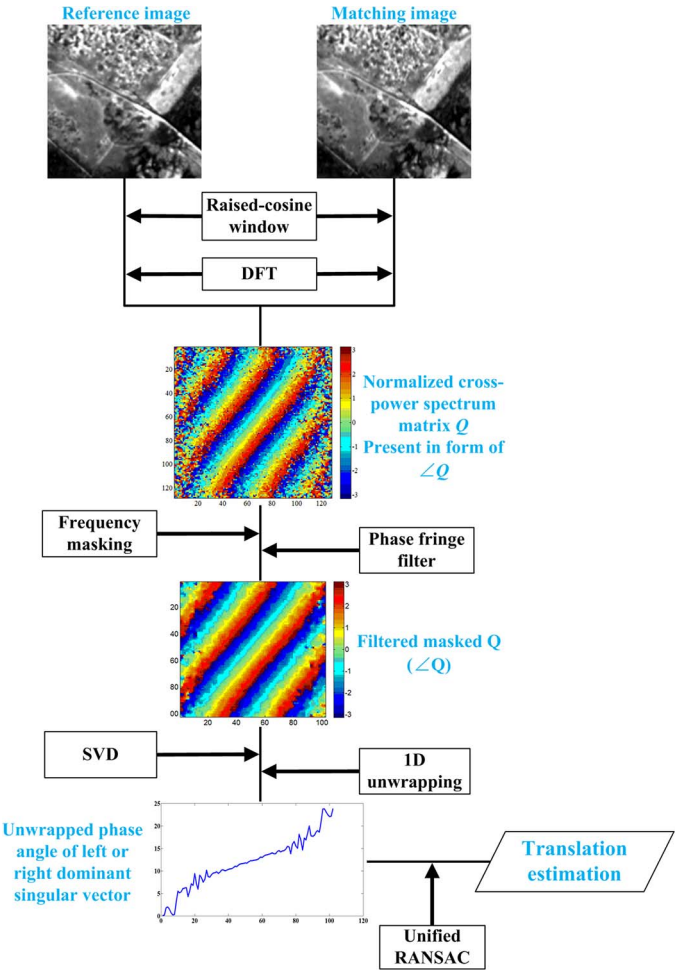

Fig. 1. Overall workflow of the proposed SVD-RANSAC method.

as suggested in [17]. 2) We compute the Fourier spectrum of the individual images using DFT and their normalized cross-power spectrum matrix, as in (3). 3) As spectral components at a high frequency are the most likely to be biased due to aliasing and noise, and most of the energy is mainly concentrated in the low-frequency components for natural images [17], [55], we should mask out useless small amounts (e.g., 10%) of frequency components at each periphery of Q. The phase fringe filter [38] that smoothes real and imaginary parts, respectively, is used to reduce the noise and weaken the influence of the corrupted phase values on the following rank-one approximation. 4) We extract the phase difference in each dimension using SVD and 1-D unwrapping. 5) The slopes (s_l, s_r) of the unwrapped phase angles of the left and right dominant singular vectors are individually identified using unified RANSAC and converted to the estimated shift parameters (x_0, y_0) as $x_0 = s_l M/2\pi$, $y_0 = -s_r N/2\pi$, where M and N denote the size of the input images.

III. Experiments Based on Simulated and Real Imagery

Both simulated and real data were used to assess the performance of the proposed SVD-RANSAC method. In the

simulated case, the images were displaced with known subpixel shifts against the ground truth. Two simulated experiments for the consideration of different error sources, namely, an aliasing experiment and a noise experiment, were carried out on synthetic image pairs from ZY-3 imagery. In the real data case, a dense matching strategy was designed to demonstrate the promising performance and feasibility of the proposed phase correlation method with the use of ZY-3 multispectral SC images with different bands.

A. Simulated Experiments for the Comparison of the Phase Correlation Methods

1) Considered Error Sources: In this paper, the two dominant error sources of phase correlation corruptions, aliasing and noise, were considered to evaluate the effectiveness and sensitivity of the subpixel phase correlation methods [17], [19], [33]. Aliasing refers to an effect that causes different signals to become indistinguishable when sampled. Aliasing occurs if images are not spatially band limited, which means that the sensor's sampling does not conform to the Nyquist–Shannon sampling theorem [37]. This is common with satellite systems designed to offer sharp images which are obtained by designing the optical cutoff frequency to be higher than the Nyquist frequency of the sensors [24]. The causes of noise are more intricate, and noise can come from both the scenes and the optical system [17]. Both cases can influence the frequency component by making the phase difference between two shifted images be more than just a linear phase shift, and consequently bias the phase correlation results. As the various subpixel phase correlation methods have different capacities to handle aliasing and noise, the robustness to these interferences is a key element to evaluating the performance of phase correlation methods.

2) Implementation Details: The results obtained from the proposed method were compared with those from six popular representative phase correlation methods: Hoge's [4], Stone's [37], Foroosh's [2], Leprince's [17], PEF [35], and Upsampling [36]. Hoge's method [4] was implemented with Stone's frequency masking [37] before the SVD step, as mentioned in the original paper. The parameter setting of the frequency masking was similar to the implementation in Stone's method [37]. Two portions were included, which mask out frequency components outside a radius from the central peak, and mask out frequency components whose corresponding spectral magnitudes of each image are less than a predefined threshold. The radius was set as 30% of the image size, and the threshold was chosen as the root-mean-square (RMS) magnitude of a 5×5 region around the center peak, multiplied by a variable of 0.01 to 0.05. Separate and consecutive 1-D unwrappings in each direction, as suggested in [34], were added to extend to more than half a pixel when necessary. As Foroosh's method [2] separately estimates the subpixel shift in each dimension, we applied the SVD approach to the normalized cross-power spectrum matrix before fitting the sinc function, to achieve accurate shift parameters, as suggested in [41]. The implementations of the PEF method [35], Leprince's method [17], and the Upsampling method [36] accorded with the original papers. We used a Gaussian function with standard deviation $\sigma = 3$ for the frequency weighting, and we set the number of reference points as $p = 3$ and the distance as $d = 5$ in the PEF method [35]. Adaptive frequency masking with the parameter $m = 0.75 - 0.9$ and four robustness iterations were used in Leprince's method [17]. The options for the two-point step size optimization algorithm were the same as the default values in the original paper. The Upsampling method was based on matrix-multiply DFT, and the upsampling factor was 1000 for all the image pairs. In the proposed method, the window size for the phase fringe filter was set relatively small, e.g., 3×3–9×9. The choice of threshold for the unified RANSAC algorithm was 0.01 to 0.03 for all the experiments in this paper. For all the methods, we used the raised-cosine window to reduce the influence of edge effects. Minor careful fine tuning was also performed to provide better estimations in each case.

3) Aliasing Experiment Steps: The purpose of this experiment was to analyze the influence of the aliasing and to compare the performance of the various subpixel phase correlation methods in the case of aliasing. The synthetic image pairs were generated with the ground truth, using the methods described in [2] and [37], which use a single real high-resolution image to generate an image pair with known subpixel shifts by low-pass filtering and downsampling. We started from a larger image and introduced a shift S of integer amounts to simulate the subpixel translation. Both the original and shifted images were then downsampled by a factor of M in each dimension, so that the relative shift became a fractional shift of S/M. To control the amount of aliasing, a low-pass filter was applied prior to downsampling [37]. In this paper, a 2-D Gaussian function with a support area size of gs and a standard deviation of σ was used as the blurring kernel. The amount of aliasing was controlled by M, gs, and σ. When M and gs are fixed, the Gaussian filter with increasing σ results in downsampled images corrupted by decreasing amounts of aliasing. The downsampling factor M was set as 10 in this experiment. The support area size gs is related to M, and was thus set as 25×25. It is worth noting that interval sampling rather than interpolation was performed, as interpolation, to some extent, amounts to a process of low-pass filtering.

The detailed steps were as follows. 1) Three subsets (1280×1280 pixels) of ZY-3 imagery were selected, as shown in Fig. 2 (upper three). 2) For each image, the subset image was translated by integer offsets ranging from 1 to 50 pixels by a step of 1 pixel in the x-direction, and fixed offset of 10 pixels in the y-direction, to achieve 50 synthetic image pairs. This offset set was adopted for expediently analyzing the integrated performance of the pixel-level and subpixel-level matching, and the pixel locking effect in the subpixel matching. 3) A Gaussian filter was applied with different standard deviations to model the camera point spread function. The amount of aliasing was controlled by changing the standard deviation σ from 1 to 5 in steps of 1. The synthetic image pairs were downsampled by a factor of 10 to obtain the fractional shifts. Fig. 3 (upper three) shows three examples of the simulated images that are generated from three raw high-resolution ZY-3 imagery. In the figure, each of the three simulated images is produced with an offset of 5 pixels in the x-direction and a standard deviation σ of 5 based

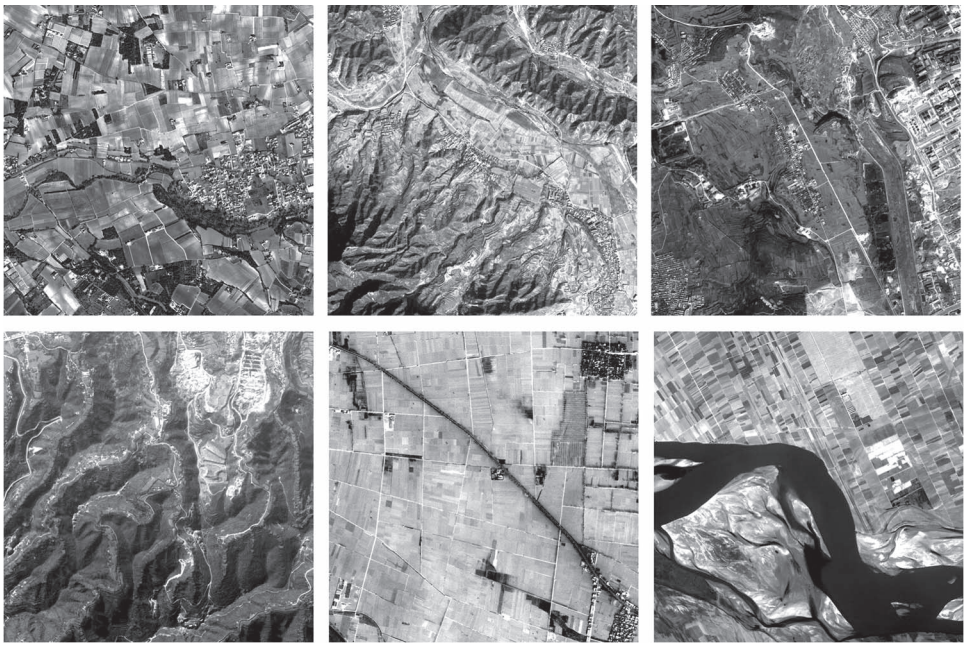

Fig. 2. Subsets of raw ZY-3 imagery used in the aliasing experiment (upper three) and noise experiment (lower three).

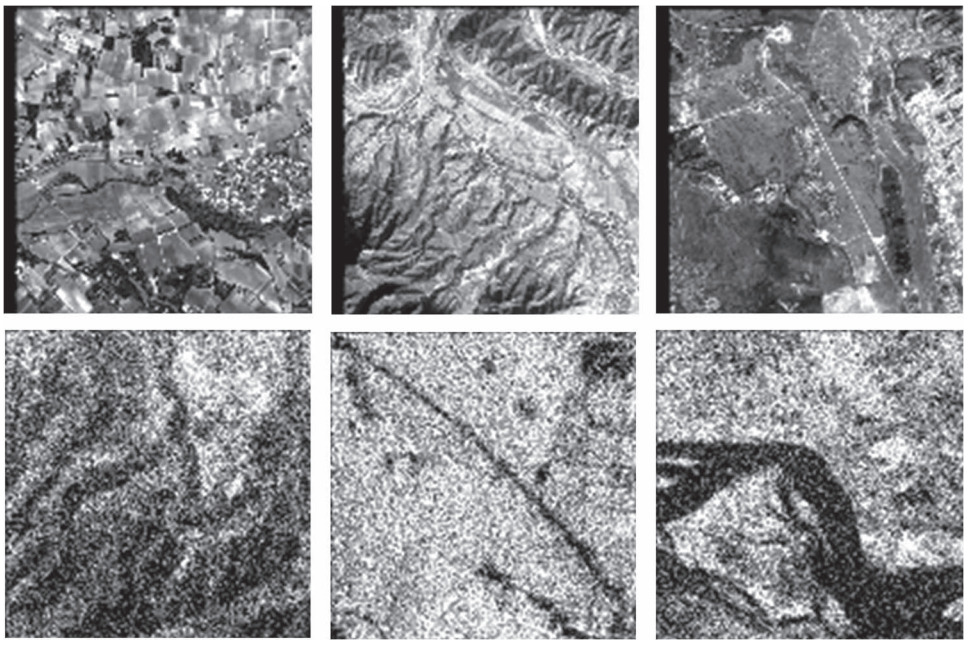

Fig. 3. Six examples of the generated simulated images in the aliasing experiment (upper three) and noise experiment (lower three).

on the raw imagery. 4) The various subpixel phase correlation methods were performed on each synthetic image pair. The results of the different methods were then computed using real displacements and were compared with each other. For each method, the performance was evaluated in terms of four measures: the mean value μ ; the RMS; the maximum value (Max); and the standard deviation σ_r of the absolute error vector between the estimates and the ground truth in the x-direction, calculated over all 150 image pairs, as a function of σ.

4) Noise Experiment Steps: In this experiment, performance assessment of the various subpixel phase correlation methods in the presence of additive noise was carried out. The synthetic image pairs with fractional shifts were generated as aforementioned. Zero-mean Gaussian white noise with different variance values was then added to the synthetic image pairs, as described in [33] and [41].

The experimental details were as follows. 1) Three subsets (1280×1280 pixels) of ZY-3 imagery were selected, as shown in Fig. 2 (lower three). Aliasing is related to the frequency spectrum of the image and is mainly present in the high-frequency regions [3], [37]. Therefore, in order to assess the separate effect of noise alone, images with relatively low variations in

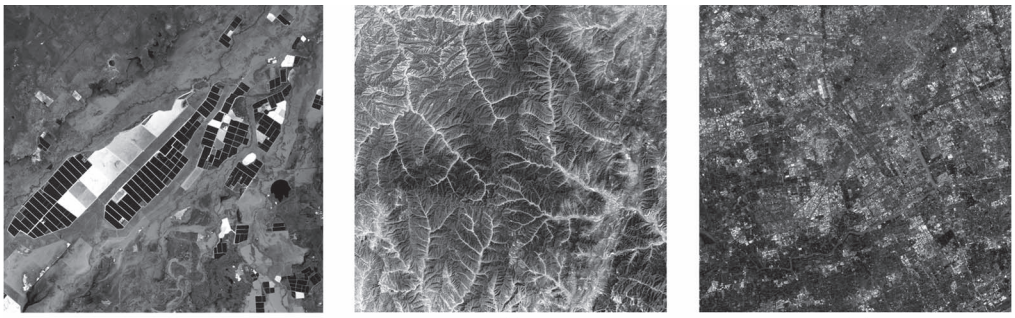

Fig. 4. ZY-3 multispectral SC images (band 1). (a) 540-783-297. (b) 5619-7-131. (c) 8332-880-147.

intensity were selected. 2) For each image, the subset image was translated by integer offsets ranging from -5 to 5 pixels by a step of 1 pixel without consideration of null displacement in each dimension. The original image and shifted images formed 100 synthetic image pairs. 3) A Gaussian filter was applied with a σ_g of 5 and a gs of 25 $\times$ 25 to blur the test image pairs and resize the image by a factor of 1/10 through bicubic interpolation, to obtain the fractional shifts. These measures also guaranteed a very small amount of aliasing. 4) Gaussian white noise was added to the synthetic image pairs. The normalized variance of noise V_n was varied from 0 to 0.05 in steps of 0.005 (the image intensity level was normalized within [0, 1]). Fig. 3 (lower three) shows three examples of the simulated images that are generated from three raw high-resolution ZY-3 imagery. In the figure, each of the three simulated images is produced with an offset of 0.5 pixels in both x- and y-directions and a normalized variance of noise V_n of 0.05 based on the raw imagery. 5) The various subpixel phase correlation methods were performed on each synthetic image pair. The correlation errors of the different methods were computed using the ground truth as $e = \sqrt{(e_x^2 + e_y^2)}$, where e_x, e_y corresponds to the error in the x- and y-directions, respectively. For each method, the same performance measures as the aliasing experiment were calculated over all 300 image pairs as a function of V_n.

B. ZY-3 Multispectral SC Images Experiment

1) Data Set and Study Area: ZY-3 is the first civilian high-resolution three-line-array stereo mapping satellite in China, and it carries three high-resolution panchromatic cameras and a multispectral camera operated in four spectral bands. The multispectral sensor has a spatial resolution of 5.8 m and consists of three charge-coupled device (CCD) units for each band. The length of overlap between the adjacent CCD units is about 195 pixels. The CCD arrays of each band are arranged in parallel on the focal plane, with certain physical distances [16].

The basic product of ZY-3 is the SC product, which is generated by applying an ideal virtual CCD array to reimage with in-flight geometric calibration parameters. The inner and exterior distortions of the SC product are considered to be resolved, and the introduced errors are negligible [56]. For the multispectral sensor, one ideal CCD array and a coarse global DEM is used to reimage for the four bands, and the virtual CCD is placed in the middle of the CCD lines of all four bands in the multispectral focal plane. Therefore, image stitching and band-to-band registration of the multispectral sensor are simultaneously achieved in the process of SC generation. The multispectral SC image is visually seamless, and the band-to-band registration accuracy has been verified to be better than 0.2 pixels [57].

In this paper, different bands of the ZY-3 multispectral SC images were used as real data to investigate the effectiveness and accuracy of the proposed phase correlation method. As shown in Fig. 4, three single-scene multispectral SC images located in different areas and acquired on different orbits, i.e., 540-783-297, 5619-7-131, and 8332-880-147, were tested.

2) Dense Matching Strategy: A pixel-to-pixel dense matching scheme based on the proposed phase correlation method was developed. The main stages are specified in the following.

1) Two different bands, e.g., band 1 and band 2 of the ZY-3 multispectral SC image are inputted. Local image enhancement via Wallis filter [58] is applied to enhance the contrast levels of the images and to reduce the luminance differences between different bands, thereby ensuring the reliability of the following dense matching.
2) Coarse correspondence between the different bands of the SC image is ensured as the band-to-band registration has been achieved in the generation of the SC image. The proposed phase correlation is directly performed between a 32 $\times$ 32 sliding subwindow pair scanning through the band 1 and band 2 images. The relative column and line shifts at every pixel, except at the boundary of the images, are determined at a subpixel accuracy, and two correlation maps are therefore generated. The subwindow size is decided with consideration of the computational efficiency and measurement accuracy. The parameter setting of the proposed phase correlation method is the same as in the simulated experiment.
3) False matches and decorrelation areas need to be removed from the correlation maps before further analysis. Three criteria are considered to eliminate the unreliable measurements. The outliers with extremely large measurements are removed at first. The featureless and low-correlation areas which have a normalized correlation coefficient (calculated in a smaller window) below a threshold

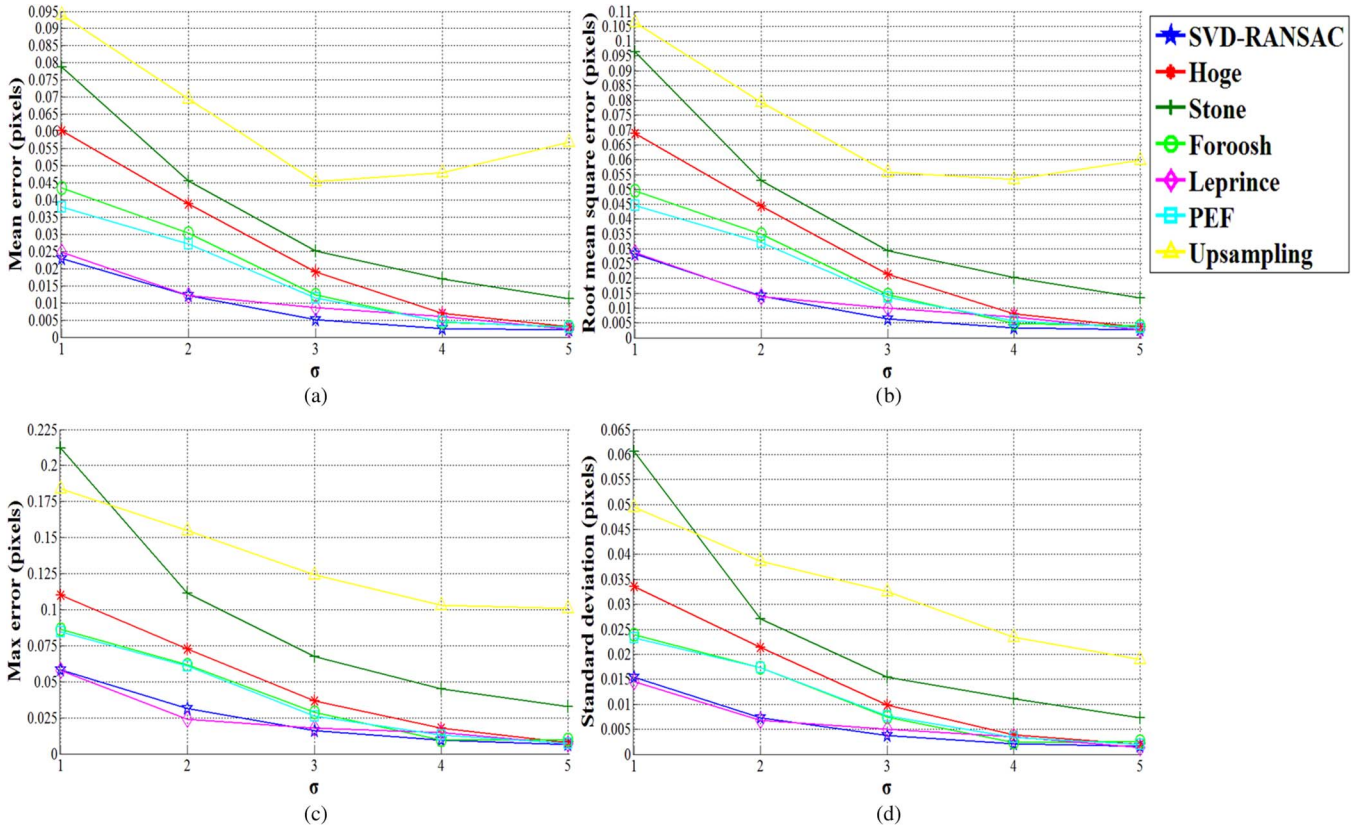

Fig. 5. (a) Mean value, (b) RMS, (c) Max value, and (d) standard deviation of the absolute error for different methods as a function of σ.

(e.g., 0.7) are also removed. Moreover, matching uncertainty measures that reflect the quality of the line fitting by the use of RANSAC during the subpixel phase correlation can be used to mask off the low correlation quality data as well.

IV. Results and Discussion

A. Simulated Data

1) Results of the Aliasing Experiment: The corresponding results of the different phase correlation methods in terms of μ, RMS, Max, and σ_r as a function of σ are shown in Fig. 5(a)–(d), respectively.

It is worth noting that the common trend for the different measures calculated by each phase correlation method reflects the overall reliability of the phase correlation method in the presence of aliasing. As can be seen, most of the methods, except for Stone's method and Upsampling, generate good results which are less than 0.005 pixels in terms of μ and RMS, with the least amount of aliasing (i.e., $\sigma = 5$), and less than 0.01 pixels when $\sigma = 4$. The proposed method as well as Leprince's method consistently yields the best results, achieving the minimum values in terms of all four measures, particularly in the case of large amounts of aliasing. The results of the μ and RMS scores are always less than 0.03 pixels. Overall, the proposed method appears slightly better than Leprince's method. With the increasing σ, Foroosh's method generates results that are comparative to PEF, which form the second-best group. Compared with the aforementioned four methods, Hoge's method performs poorly and is more susceptible to the aliasing effect. Among all the methods, Stone's method and Upsampling produce the worst results. To some extent, the complicated but indispensable 2-D phase unwrapping could bias the estimates of Stone's method. It is, however, surprising that the performance of Upsampling in terms of μ and RMS becomes worse when $\sigma > 3$. The similar case can be seen in [37] and [41].

2) Pixel Locking Effect Analysis: In the aliasing experiment, synthetic image pairs with relative offsets ranging from 0.1 to 5 pixels in the x-direction and a fixed offset in the y-direction were generated. An error distribution analysis was further conducted to investigate the performance of the different phase correlation methods under various offsets in one dimension. When the horizontal offset varied at a certain value, the mean value of the correlation errors among the three test images at each value of σ for each method was calculated. Fig. 6 illustrates a comparison of the mean errors obtained from Hoge's method and SVD-RANSAC with the real horizontal offsets. The scales of the two curves were set the same, for the purpose of a straightforward comparison.

As can be seen, the errors generated when the given offset exceeds one pixel are similar to the ones generated under a subpixel offset for both Hoge's method and SVD-RANSAC. This reflects the ability of Hoge's method to simultaneously estimate integral and fractional shifts as well as SVD-RANSAC. In addition, it is noteworthy that the two sets of curves present a systematic trend, which indicates that the estimates contain

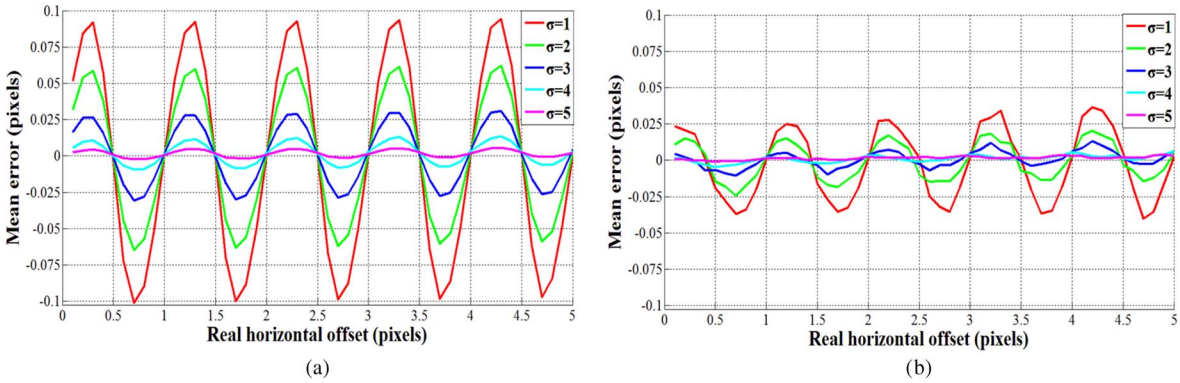

Fig. 6. Mean errors under various offsets at each value of σ. (a) Results that are obtained from Hoge's method. (b) Results that are obtained from SVD-RANSAC.

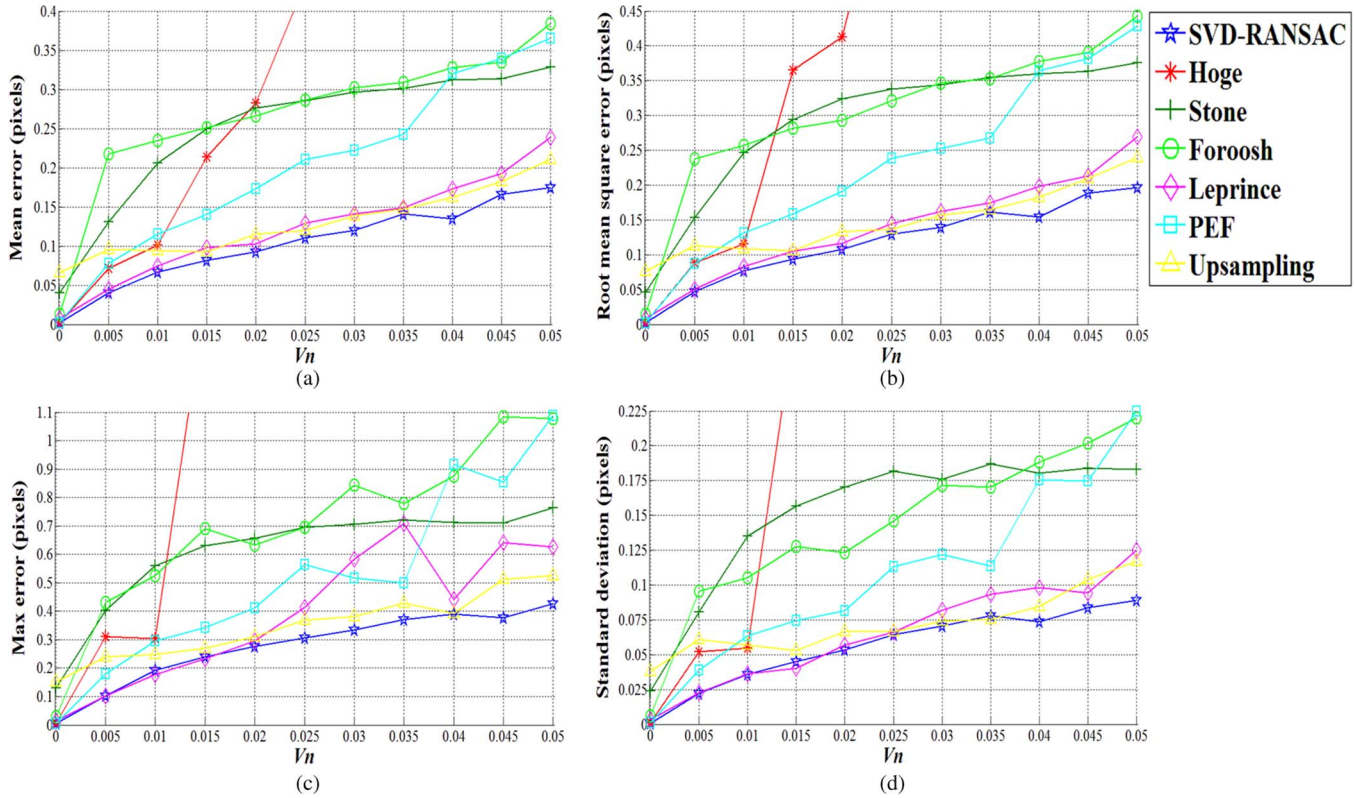

Fig. 7. (a) Mean value, (b) RMS, (c) Max value, and (d) standard deviation of the absolute error for the different methods as a function of V_n. The results of Hoge's method were truncated for a better visualization.

a systematic error, depending on the image characteristics and correlation algorithms. This systematic error, the so-called pixel locking effect, has been reported in the field of particle image velocimetry and computer vision [59]–[61]. The pixel locking effect represents the phenomenon by which subpixel estimates bias toward integral pixel positions. As shown in Fig. 6, the pixel locking effect also exists in the subpixel phase correlation. This results in a periodic error with a pixel interval that is relatively small for shifts at approximate multiples of 0.5 pixels and at integral pixel locations, and is antisymmetric with respect to these locations. The obvious fact that the degree of pixel locking effect decreases with an increasing value of σ demonstrates that aliasing is one of the crucial causes and influence factors of the pixel locking effect. A large amount of aliasing, indicating incomplete sampling with the sensor, worsens the pixel locking effect. It is found that the proposed method weakened the pixel locking effect at each value of σ, due to the improved robustness to aliasing, as compared with Hoge's method. However, a certain degree of pixel locking effect still remains, particularly in the case of a large amount of aliasing. The specific causes of the pixel locking effect are complicated, and need further research.

3) Results of the Noise Experiment: The performances of the different phase correlation methods as V_n was varied between [0, 0.05] are illustrated in Fig. 7(a)–(d) in terms of μ, RMS, Max, and σ_r. Regarding the measurements in Fig. 7, Hoge's method appears to be the most sensitive to noise and achieves the worst results. In particular, in the case of increasing

normalized variance of noise, the correct pixel accuracy is not always given. Conversely, it is again found that the results generated by the proposed method show the most precise overall measurements and the most stable performance among all the test methods. As far as μ and RMS are concerned, the proposed SVD-RANSAC outperforms the original Hoge's method by at least three times in the case of high-level noise (i.e., $V_n \geq 0.02$). The results of Hoge's method were truncated for better visualization, as its significa\footnote{text}nt errors cover a large data range.

A simple visual inspection of Fig. 7 shows that Leprince's method again achieves the second place in terms of the performance over the entire set of test conditions. Although Upsampling generates poor estimates in the noise-free case, it features the most robust performance as the noise variance increases, and it defeats the other methods, except for the proposed method, for high-level noise. The reason for this could be that only a small neighborhood is used for the final subpixel estimation. The Stone's, Foroosh's, and PEF methods provide less accurate correlation results compared with the top three methods, but they appear to be less sensitive to noise than Hoge's method. Moreover, for every performance measure, the performance of PEF linearly varies with respect to the noise variance, while the variation curve of Stone's method represents a quadratic shape. The error of Foroosh's method rapidly increases once additive noise is embedded, and it gradually increases with more serious noise.

4) Discussion on Superiority in Robustness and Relative Computational Efficiency: As aforementioned, the superiority of the proposed method over the other state-of-the-art phase correlation methods was confirmed according to the minimum overall results of four performance measures under a variety of test conditions, considering aliasing and noise. The proposed method inherits the advantages of Hoge's methods and avoids its nonignorable shortfalls. Therefore, our method generates estimates with both high accuracy and powerful robustness against diverse forms of interference, including aliasing and noise. In contrast, the other phase correlation methods, such as Stone's, Foroosh's, Leprince's, PEF, and Upsampling, were shown to be more sensitive to aliasing and noise. By synthesizing the results of the two simulated experiments, Leprince's method provides the second-best performance, and it benefits from its correlation scheme, including nonlinear optimization, adaptive frequency masking, and robustness iteration [17].

As aforementioned, aliasing and noise are the two main error sources. Some researchers have adopted certain measures to explicitly deal with aliasing and noise. As the high frequencies are the most likely to be degraded due to aliasing and noise, Kim and Su [62] relied on the low-frequency components of the transforms to eliminate the aliasing effects. Nagashima *et al.* [35] also applied a low-pass-type Gaussian weighting function on the normalized cross-spectrum matrix in the frequency domain to eliminate the high-frequency components with a low reliability. Stone *et al.* [37] and Leprince *et al.* [17] shared the consensus that it is not sufficient to merely mask out high frequencies, and the frequencies that are most likely to be corrupted are those with a small spectral magnitude or cross-spectrum magnitude. In [37], the spectral components of each image with a magnitude lower than a threshold were further masked out. In [17], the adaptive frequency masking was defined by retaining only the components where the magnitude of the normalized log spectrum exceeds the mean magnitude of the entire normalized log spectrum multiplied by a parameter threshold. However, "small" is a relative concept, which means that the threshold in frequency masking is arbitrarily chosen. Thus, the parameter setting largely determines the behavior of the method, as indicated by the aforementioned experiment in this paper. In the proposed method, similar to the ideas in [38] and [63], we adopt ideas of robust estimation (i.e., unified RANSAC in this case) with a known model (i.e., linear model) to exclude the corruption due to aliasing and noise. The use of the RANSAC algorithm is the key factor in our method, and it equips the method with accuracy and robustness. Therefore, the performance of the proposed method is mainly controlled by the predefined threshold in RANSAC. Unified RANSAC is applied in this case to ensure the stability of the results and to minimize the sensitivity to the threshold, as demonstrated by the experimental data. In addition, as mentioned in [64], larger interpolation errors for the low-frequency components are induced by Cartesian-to-log-polar conversion in the Fourier–Mellin transformation, as log-polar representation is extremely dense in the low frequency. Additional sources of aliasing artifacts are also induced by rotations and scalings of the images [65]. The proposed method with a high accuracy and robustness is therefore suitable to be extended to accurate image registration for the estimation of translation, rotation, and scale.

A comparison of computational time and mean absolute error (μ) of the proposed SVD-RANSAC method and the other six ones in the aliasing experiment is provided in Table I. The total computational counts with respect to each method are $50 \times 3 \times 5 = 750$ times. From the table, we can see that 1) the interpolation methods, i.e., Foroosh's method and PEF, show considerably high computational efficiency. However, their accuracy and robustness are inferior to both Leprince's method and the proposed SVD-RANSAC method. 2) The computational time of Upsampling method is determined by the upsampling factor (i.e., 1000 in this study) and is dominated by the initial estimate using the conventional fast FT upsampling method [36]. 3) The computational time of Stone's method is restricted by the time-consuming step of 2-D phase unwrapping. By contrast, Hoge's method uses SVD to find the dominant rank-one approximation of the normalized cross-power spectrum matrix and thus it merely needs 1-D phase unwrapping. As a result, its computational time is greatly reduced. 4) The proposed method, which combines Hoge's method with unified RANSAC to improve the accuracy and robustness, spends a little more computational time than the first three methods. However, it spends much less computational time than Upsampling, Stone's and Leprince's methods.

All these seven methods are programmed in MATLAB on a computer with 3.0-GHz CPU, and the details of the implementation are described in Section III-A2. As shown in Table I, the computational time of the proposed SVD-RANSAC method is less than that of Leprince's method, and our proposed method yields the best accuracy in terms of mean absolute errors.

TABLE I
COMPARISON OF COMPUTATIONAL TIME (SECOND) AND MEAN ABSOLUTE ERROR μ (PIXEL) OF THE PROPOSED SVD-RANSAC METHOD AND THE OTHER SIX ONES IN THE ALIASING EXPERIMENT

Method	Foroosh	PEF	Hoge	Upsampling	Stone	Leprince	Proposed SVD-RANSAC
Total time / second	9.491	6.019	12.810	190.919	120.265	157.093	30.188
μ / pixel (σ =1)	0.0435	0.0380	0.0603	0.0940	0.0787	0.0248	0.0230
μ / pixel (σ =2)	0.0303	0.0271	0.0390	0.0695	0.0456	0.0121	0.0122
μ / pixel (σ =3)	0.0124	0.0113	0.0191	0.0453	0.0251	0.0087	0.0052
μ / pixel (σ =4)	0.0043	0.0046	0.0070	0.0479	0.0170	0.0060	0.0025
μ / pixel (σ =5)	0.0031	0.0028	0.0031	0.0568	0.0113	0.0023	0.0021

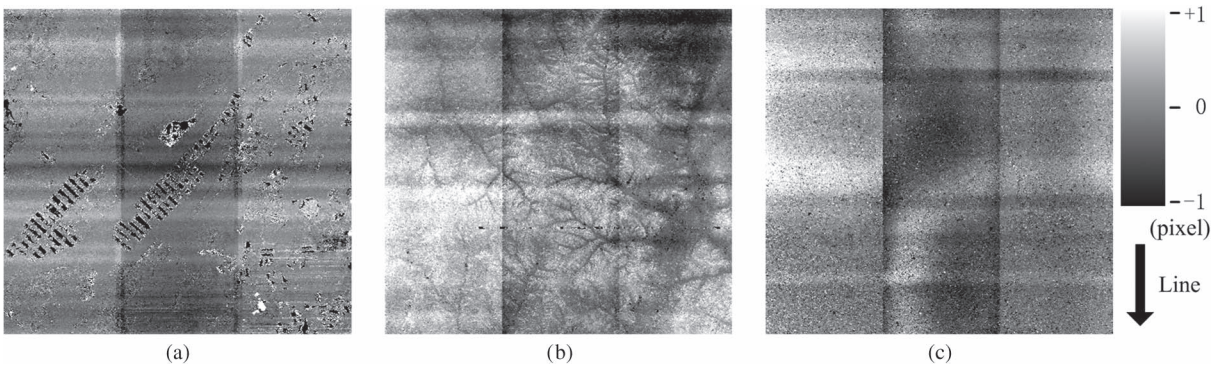

Fig. 8. Correlation maps of band 1 relative to band 2 in the cross-track direction. (a) 540-783-297. (b) 5619-7-131. (c) 8332-880-147.

Therefore, the proposed SVD-RANSAC method is reliable and acceptable in view of both accuracy and computational efficiency.

B. Dense Matching Using ZY-3 Multispectral SC Images

1) Results of the Dense Matching: In order to demonstrate the effectiveness and feasibility of the proposed phase correlation method in real application, band 1 and band 2 of the ZY-3 multispectral SC images described in Section III-B2 were used as experimental data. We correlated the two images by means of the designed dense matching scheme (Section III-B3) pixel by pixel and obtained two filtered correlation maps. In this case, the correlation map in the cross-track direction is presented in Fig. 8 to analyze the image stitching situation of the multispectral SC image. The gray variations in the figures represent the disparities between band 1 and band 2 in the cross-track direction. As the band-to-band registration of the multispectral sensor has been achieved in the process of SC generation, measurements over 1 pixel were considered as outliers and removed from the correlation maps. Subsequently, the filtered correlation map was averaged in each column, as shown in Fig. 9, in order to reflect the integral condition of the whole virtual CCD array.

2) Results Analysis: As shown in Fig. 8, decorrelation areas, which are presented as the dark shade in the correlation maps, were identified and filtered out. For different bands of the multispectral image with a short imaging time separation and very small view angle difference, the correlation is mainly lost in the featureless regions with constant radiometry and lack of texture, such as the farmland and mountain shadow. With the exception of these regions, few decorrelation areas were generated, which confirms the effectiveness of the proposed method.

From Fig. 8, it can be observed that the three sub-CCD units of the multispectral sensor present different overall disparity levels. Differences in the disparities with respect to two adjacent units are explicitly identifiable in the correlation maps. This phenomenon results from residual errors in the image stitching. Fig. 9 illustrates this case more clearly. The apparent sharp discontinuities are found exactly in the overlapping areas between adjacent units, and they represent the corresponding

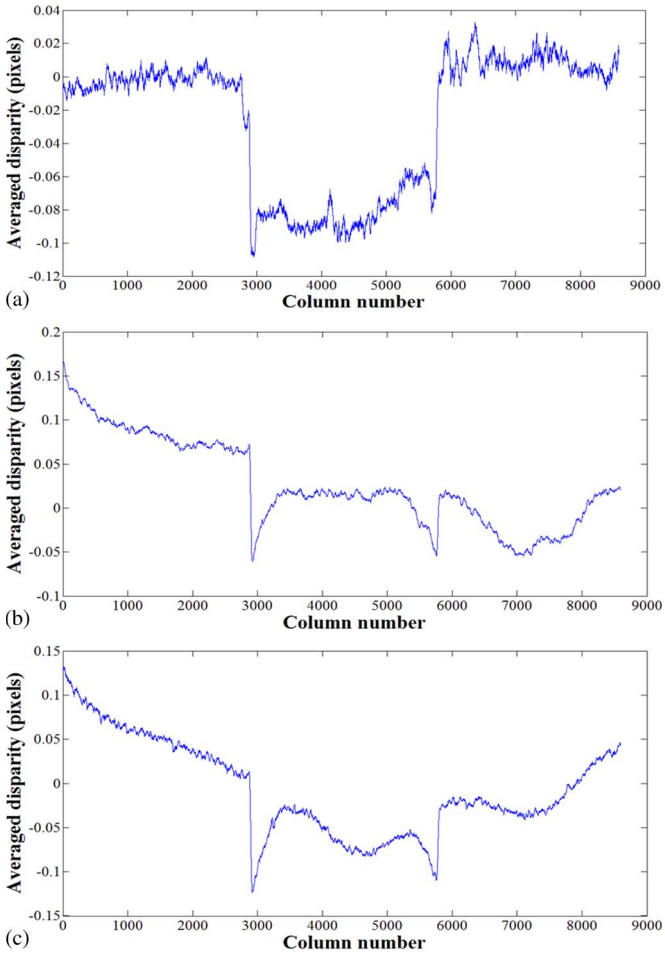

Fig. 9. Averaged disparity at each column of band 1 relative to band 2 in the cross-track direction. (a) 540-783-297. (b) 5619-7-131. (c) 8332-880-147.

locations of the seams of the image stitching. In contrast, the averaged disparity within each sub-CCD unit appears relatively stable, without obvious high-frequency noise. Moreover, the RMS of the disparity at each pixel, excluding outliers, is 0.098, 0.091, and 0.086 pixels for the three multispectral SC images, respectively. The accuracies are all below 0.1 pixels, which is in agreement with the band-to-band registration accuracy reported in [16] and [57].

In addition, some wavy patterns with light and dark variations in the line direction appear in the correlation maps. These unexpected strips correspond to unmodeled satellite attitude jitter during the image acquisition, which can be further analyzed and compensated using raw data without geometric processing [25], [26].

V. Conclusion

In this paper, a novel robust subpixel phase correlation method based on singular value decomposition and the unified RANSAC algorithm has proposed. Precise and reliable subpixel estimations can be directly achieved in the Fourier domain. We attribute the remarkable merits of the proposed method to the stability and robustness of the unified RANSAC algorithm. These merits are conducive to improving the performance of further remote sensing applications based on subpixel phase correlation. The performance of the proposed method was demonstrated with experiments using both simulated and real data. The results obtained under a variety of test conditions confirm the superiority and feasibility of the proposed method. Several conclusions can be drawn, as follows.

1) In the simulated data experiments, the proposed method consistently generated the best results in both the aliasing and noise experiments, compared with the other six current prominent phase correlation methods. Particularly in the noise experiment, Hoge's method provided the worst performance among the test methods. It was also proven that improvement based on unified RANSAC plays a significant role in enhancing the robustness and stability.
2) The degree of pixel locking effect increases with increasing amounts of aliasing, which is visibly weakened by the proposed method, due to the improved robustness to aliasing over Hoge's method, but it cannot be completely eliminated.
3) In the real data experiments, different bands of the ZY-3 multispectral SC images were coregistered using the designed pixel-to-pixel dense matching scheme. The results showed that seams of the image stitching between the different sub-CCD units could be clearly identified. Furthermore, the calculated RMS of the disparity between different bands accords with the published band-to-band registration accuracy. Again, the results confirm the effectiveness and reliability of the proposed method.

This paper focuses on image registration for translation estimation. In future work, the proposed method could be considered as suitable to extend to accurate translation, rotation, and scale-invariant image registration by means of Fourier–Mellin transformation, on the basis of the Fourier scaling property and the Fourier rotational property [66], [67]. In addition, a detailed investigation of the systematic correlation errors, including the pixel locking effect, will be the future focus of our research.

Acknowledgment

The authors would like to the editors and the anonymous reviewers for their valuable comments on the improvement of this paper.

References

[1] T. Heid and A. Kääb, "Evaluation of existing image matching methods for deriving glacier surface displacements globally from optical satellite imagery," *Remote Sens. Environ.*, vol. 118, pp. 339–355, Mar. 2012.
[2] 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, Mar. 2002.
[3] P. Vandewalle, S. Sü, and M. Vetterli, "A frequency domain approach to registration of aliased images with application to super-resolution," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 1–14, 2006.
[4] W. S. Hoge, "A subspace identification extension to the phase correlation method [MRI application]," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 277–280, Feb. 2003.
[5] H. T. Ho and R. Goecke, "Optical flow estimation using Fourier Mellin transform," in *Proc. CVPR*, 2008, pp. 1–8.
[6] S. Kumar, H. Azartash, M. Biswas, and T. Nguyen, "Real-time affine global motion estimation using phase correlation and its application for digital image stabilization," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3406–3418, Dec. 2011.

[7] W. Pan, K. Qin, and Y. Chen, "An adaptable-multilayer fractional Fourier transform approach for image registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 3, pp. 400–414, Mar. 2009.

[8] K. Ito, H. Nakajima, K. Kobayashi, and T. Higuchi, "A fingerprint matching algorithm using phase-only correlation," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. 87, no. 3, pp. 682–691, 2004.

[9] M. A. Muquit and T. Shibahara, "A high-accuracy passive 3D measurement system using phase-based image matching," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. 89, no. 3, pp. 686–697, 2006.

[10] C. Dai, Y. Zheng, and X. Li, "Accurate video alignment using phase correlation," *IEEE Signal Process. Lett.*, vol. 13, no. 12, pp. 737–740, Dec. 2006.

[11] M. Paul, W. Lin, C. T. Lau, and B.-s. Lee, "Direct intermode selection for H. 264 video coding using phase correlation," *IEEE Trans. Image Process.*, vol. 20, no. 2, pp. 461–473, Feb. 2011.

[12] J. Liu and H. Yan, "Phase correlation pixel-to-pixel image coregistration based on optical flow and median shift propagation," *Int. J. Remote Sens.*, vol. 29, no. 20, pp. 5943–5956, Oct. 2008.

[13] G. L. K. Morgan, J. G. Liu, and H. Yan, "Precise subpixel disparity measurement from very narrow baseline stereo," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3424–3433, Sep. 2010.

[14] T. Arai and A. Iwasaki, "Fine image matching for narrow baseline stereo-vision," in *Proc. IEEE IGARSS*, Munich, Germany, 2012, pp. 2336–2339.

[15] S. Leprince, P. Musé, and J.-P. Avouac, "In-flight CCD distortion calibration for pushbroom satellites based on subpixel correlation," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 9, pp. 2675–2683, Sep. 2008.

[16] Y. Jiang *et al.*, "Geometric calibration and accuracy assessment of ZiYuan-3 multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4161–4172, Jul. 2014.

[17] S. Leprince, S. Barbot, F. Ayoub, and J.-P. Avouac, "Automatic and precise orthorectification, coregistration, and subpixel correlation of satellite images, application to ground deformation measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1529–1558, Jun. 2007.

[18] N. Van Puymbroeck, R. Michel, R. Binet, J.-P. Avouac, and J. Taboury, "Measuring earthquakes from optical satellite images," *Appl. Opt.*, vol. 39, no. 20, pp. 3486–3494, Jul. 2000.

[19] P. J. Gonzalez, M. Chini, S. Stramondo, and J. Fernández, "Coseismic horizontal offsets and fault-trace mapping using phase correlation of IRS satellite images: The 1999 Izmit (Turkey) earthquake," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2242–2250, May 2010.

[20] D. Scherler, S. Leprince, and M. R. Strecker, "Glacier-surface velocities in alpine terrain from optical satellite imagery—Accuracy improvement and quality assessment," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3806–3819, Oct. 2008.

[21] M. Necsoiu *et al.*, "Monitoring migration rates of an active subarctic dune field using optical imagery," *Remote Sens. Environ.*, vol. 113, no. 11, pp. 2441–2447, Nov. 2009.

[22] N. Bridges *et al.*, "Earth-like sand fluxes on Mars," *Nature*, vol. 485, no. 7398, pp. 339–342, 2012.

[23] E. Berthier *et al.*, "Surface motion of mountain glaciers derived from satellite optical imagery," *Remote Sens. Environ.*, vol. 95, no. 1, pp. 14–28, Mar. 2005.

[24] S. Leprince, "Monitoring Earth surface dynamics with optical imagery," Ph.D. dissertation, California Institute of Technology, Pasadena, CA, USA, 2008.

[25] Y. Teshima and A. Iwasaki, "Correction of attitude fluctuation of Terra spacecraft using ASTER/SWIR imagery with parallax observation," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 222–227, Jan. 2008.

[26] X. Tong *et al.*, "Framework of jitter detection and compensation for high resolution satellites," *Remote Sens.*, vol. 6, no. 5, pp. 3944–3964, 2014.

[27] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.

[28] S. Baker *et al.*, "A database and evaluation methodology for optical flow," *Int. J. Comput. Vis.*, vol. 92, no. 1, pp. 1–31, Mar. 2011.

[29] C. Kuglin, "The phase correlation image alignment method," in *Proc. IEEE Conf. Cybern. Soc.*, 1975, pp. 163–165.

[30] Q. Tian and M. N. Huhns, "Algorithms for subpixel registration," *Comput. Vis., Graph. Image Process.*, vol. 35, no. 2, pp. 220–233, Aug. 1986.

[31] I. E. Abdou, "Practical approach to the registration of multiple frames of video images," in *Proc. SPIE Conf. Vis. Comm. Image Process.*, 1999, pp. 371–382.

[32] V. Argyriou and T. Vlachos, "A Study of Sub-pixel Motion Estimation using Phase Correlation," in *Proc. BMVC*, 2006, pp. 387–396.

[33] J. Ren, J. Jiang, and T. Vlachos, "High-accuracy sub-pixel motion estimation from noisy images in Fourier domain," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1379–1384, May 2010.

[34] H. Foroosh and M. Balci, "Sub-pixel registration and estimation of local shifts directly in the Fourier domain," in *Proc. Int. Conf. Image Process.*, 2004, pp. 1915–1918.

[35] S. Nagashima, T. Aoki, T. Higuchi, and K. Kobayashi, "A subpixel image matching technique using phase-only correlation," in *Proc. IEEE 2006 Int. Symp. Intell. Signal Process.Commun. Syst.*, 2006, pp. 701–704.

[36] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Opt. Lett.*, vol. 33, no. 2, pp. 156–158, 2008.

[37] H. S. Stone, M. T. Orchard, E.-C. Chang, and S. A. Martucci, "A fast direct Fourier-based algorithm for subpixel registration of images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 10, pp. 2235–2243, Oct. 2001.

[38] J. G. Liu and H. Yan, "Robust phase correlation methods for sub-pixel feature matching," in *Proc. 1st Annu. Conf. Syst. Eng. Auton. Syst., Defence Technol. Centre*, Edinburgh, U.K., 2006, p. A13.

[39] Y. Keller and A. Averbuch, "A projection-based extension to phase correlation image alignment," *Signal Process.*, vol. 87, no. 1, pp. 124–133, Jan. 2007.

[40] R. Mumtaz, P. Palmer, and M. Waqar, "Georeferencing of UK DMC stereo-images without ground control points by exploiting geometric distortions," *Int. J. Remote Sens.*, vol. 35, no. 6, pp. 2136–2169, 2014.

[41] G. Tzimiropoulos, V. Argyriou, and T. Stathaki, "Subpixel registration with gradient correlation," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1761–1767, Jun. 2011.

[42] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

[43] O. Chum and J. Matas, "Optimal randomized RANSAC," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 8, pp. 1472–1482, Aug. 2008.

[44] H. Wang and D. Suter, "MDPE: A very robust estimator for model fitting and range image segmentation," *Int. J. Comput. Vis.*, vol. 59, no. 2, pp. 139–166, Sep. 2004.

[45] R. Raguram, O. Chum, M. Pollefeys, J. Matas, and J. Frahm, "USAC: A Universal Framework for Random Sample Consensus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 2022–2038, Aug. 2013.

[46] T. Kim and Y.-J. Im, "Automatic satellite image registration by combination of matching and random sample consensus," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 5, pp. 1111–1117, May 2003.

[47] A. Wong and D. A. Clausi, "ARRSI: Automatic registration of remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1483–1493, May 2007.

[48] P. H. S. Torr, "Motion segmentation and outlier detection," Ph.D. dissertation, Univ. Oxford, Oxford, U.K., 1995.

[49] D. Scaramuzza, "1-point-ransac structure from motion for vehicle-mounted cameras by exploiting non-holonomic constraints," *Int. J. Comput. Vis.*, vol. 95, no. 1, pp. 74–85, 2011.

[50] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[51] S. Choi, T. Kim, and W. Yu, "Performance evaluation of RANSAC family," in *Proc. BMVC*, 2009, pp. 1–12.

[52] P. H. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Comput. Vis. Image Understand.*, vol. 78, no. 1, pp. 138–156, Apr. 2000.

[53] Q. H. Tran, T.-J. Chin, W. Chojnacki, and D. Suter, "Sampling minimal subsets with large spans for robust estimation," *Int. J. Comput. Vis.*, vol. 106, no. 1, pp. 93–112, Jan. 2014.

[54] O. Chum, J. Matas, and J. Kittler, "Locally optimized RANSAC," in *Proc. Ann. Symp. German Assoc. Pattern Recognit.*, 2003, pp. 236–243.

[55] A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *Int. J. Comput. Vis.*, vol. 42, no. 3, pp. 145–175, May 2001.

[56] H. Pan *et al.*, "Basic products of the ZiYuan-3 satellite and accuracy evaluation," *Photogramm. Eng. Remote Sens.*, vol. 79, no. 12, pp. 1131–1145, Dec. 2013.

[57] X. Tang, P. Zhou, G. Zhang, and X. Wang, "Research on a Production Method of Sensor Corrected Products for ZY-3 Satellite," *Geomat. Inf. Sci. Wuhan Univ.*, vol. 39, no. 3, pp. 287–294, 2014.

[58] I. Jazayeri and C. S. Fraser, "Interest operators for feature-based matching in close range photogrammetry," *Photogramm. Rec.*, vol. 25, no. 129, pp. 24–41, Mar. 2010.

[59] J. Westerweel, "Fundamentals of digital particle image velocimetry," *Meas. Sci. Technol.*, vol. 8, no. 12, p. 1379, Dec. 1997.

[60] M. Shimizu and M. Okutomi, "Sub-pixel estimation error cancellation on area-based matching," *Int. J. Comput. Vis.*, vol. 63, no. 3, pp. 207–224, Jul. 2005.

[61] M. R. Cholemari, "Modeling and correction of peak-locking in digital PIV," *Exp. Fluids*, vol. 42, no. 6, pp. 913–922, Jun. 2007.
[62] S. Kim and W.-Y. Su, "Subpixel accuracy image registration by spectrum cancellation," in *Proc. IEEE ICASSP*, 1993, pp. 153–156.
[63] M. Balci and H. Foroosh, "Subpixel estimation of shifts directly in the Fourier domain," *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 1965–1972, Jul. 2006.
[64] G. Tzimiropoulos, V. Argyriou, S. Zafeiriou, and T. Stathaki, "Robust FFT-based scale-invariant image registration with image gradients," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 10, pp. 1899–1906, Oct. 2010.
[65] H. S. Stone, B. Tao, and M. McGuire, "Analysis of image registration noise due to rotationally dependent aliasing," *J. Vis. Comm. Image Representation*, vol. 14, no. 2, pp. 114–135, 2003.
[66] Q. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 12, pp. 1156–1168, Dec. 1994.
[67] B. S. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE Trans. Image Process.*, vol. 5, no. 8, pp. 1266–1271, Aug. 1996.

Xiaohua Tong received the Ph.D. degree from Tongji University, Shanghai, China, in 1999.

Between 2001 and 2003, he was a Postdoctoral Researcher with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China. He was a Research Fellow with Hong Kong Polytechnic University, Kowloon, Hong Kong, in 2006, and a Visiting Scholar with the University of California, Santa Barbara, CA, USA, between 2008 and 2009. His current research interests include remote sensing, geographic information system, quality and trust in spatial data, and image processing for high-resolution and hyperspectral images.

Dr. Tong serves as the Vice Chair of the commission on data quality of International Cartographical Association and a Cochair of working group on spatial data quality of the International Society for Photogrammetry and Remote Sensing.

Zhen Ye was born in 1989. He received the B.E. degree from Tongji University, Shanghai, China, in 2011. He is currently working toward the Ph.D. degree in the College of Surveying and Geo-informatics, Tongji University.

His research interests include photogrammetry and remote sensing, high-accuracy image matching, and high-resolution satellite image processing.

Yusheng Xu was born in 1989. He received the B.E. and M.E. degrees from Tongji University, Shanghai, China, in 2011 and 2014, respectively. He is currently working toward the Ph.D. degree in the Department of Photogrammetry and Remote Sensing, Technical University of Munich, Munich, Germany.

His research interests are in image processing, stereo matching, space photogrammetry, point clouds processing, and 3-D reconstruction.

Shijie Liu was born in 1982. He received the B.S., M.S., and Ph.D. degrees from Tongji University, Shanghai, China, in 2005, 2008, and 2012, respectively.

He is currently a Lecturer of surveying and mapping science and technology with the College of Surveying and Geoinformatics, Tongji University. His research interests are in geometric exploitation of high-resolution remote sensing imagery and its application.

Lingyun Li received the B.S. degree in geographical information system from Yunnan University, Yunnan, China in 2008. He is currently working toward the Ph.D. degree in the College of Surveying and Geo-informatics, Tongji University, Shanghai, China.

From 2012 to 2013, he was a Visiting Student with the Department of Planetary Geodesy, Technical University of Berlin, Berlin, Germany. His research mainly focuses on planetary mapping, photogrammetry and remote sensing, and geographic information systems.

Huan Xie received the B.S. degree in surveying engineering, the M.S. and Ph.D. degrees in cartography and geoinfomation from Tongji University, Shanghai, China, in 2003, 2006, and 2009, respectively.

From 2007 to 2008, she was a Visiting Student with the Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Hannover, Germany. She is currently an Associate Professor with the College of Surveying and Geo-Informatics, Tongji University. Her research interests include hyperspectral remote sensing and polar remote sensing.

Tianpeng Li was born in 1989. He received the B.E. degree from the China University of Geosciences, Wuhan, China, in 2012. He is currently working toward the master's degree in the College of Surveying and Geo-informatics, Tongji University, Shanghai, China.

His research interests include photogrammetry and remote sensing, high-resolution satellite image processing and DEM generation.