

## Research Article

# A Registration Scheme for Multispectral Systems Using Phase Correlation and Scale Invariant Feature Matching

Hanlun Li,<sup>1</sup> Aiwu Zhang,<sup>1</sup> and Shaoxing Hu<sup>2</sup>

<sup>1</sup>Key Laboratory of 3D Information Acquisition and Application of Ministry of Education, Capital Normal University, Beijing 100048, China

<sup>2</sup>School of Mechanical Engineering & Automation, Beihang University, Beijing 100083, China

Correspondence should be addressed to Aiwu Zhang; zhangaw98@163.com

Received 3 July 2015; Revised 20 November 2015; Accepted 24 November 2015

Academic Editor: Hairong Qi

Copyright © 2016 Hanlun Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the past few years, many multispectral systems which consist of several identical monochrome cameras equipped with different bandpass filters have been developed. However, due to the significant difference in the intensity between different band images, image registration becomes very difficult. Considering the common structural characteristic of the multispectral systems, this paper proposes an effective method for registering different band images. First we use the phase correlation method to calculate the parameters of a coarse-offset relationship between different band images. Then we use the scale invariant feature transform (SIFT) to detect the feature points. For every feature point in a reference image, we can use the coarse-offset parameters to predict the location of its matching point. We only need to compare the feature point in the reference image with the several near feature points from the predicted location instead of the feature points all over the input image. Our experiments show that this method does not only avoid false matches and increase correct matches, but also solve the matching problem between an infrared band image and a visible band image in cases lacking man-made objects.

## 1. Introduction

Many multispectral systems were developed in recent years. For example, Oppelt and Mauser [1] introduced an imaging system. Once this imaging system has been integrated, it is difficult to change the bandpass filters to acquire the band images at other wavelengths. Gorsevski and Gessler [2] designed an airborne mapping system that can provide valuable experiential learning opportunities for students. The multispectral system is only one subsystem of the complex airborne mapping system, and they did not introduce the registration method. Yang et al. [3] developed an airborne multispectral system which consists of four identical monochrome cameras equipped with four bandpass filters, and they used polynomial transformation models to register different band images. However, they did not introduce the method to generate matching points. Wu et al. [4] also developed an airborne multispectral system with four monochrome cameras and four bandpass filters, and they used the homography

to produce multispectral images. As with Yang et al., Wu et al. did not introduce how to generate matching points. Most of these papers focus on hardware design and data synchronization acquisition but fail to introduce registration methods in detail. We urgently need an effective registration method for the multispectral systems. In light of this, an airborne multispectral system which consists of four identical monochrome cameras equipped with four bandpass filters was developed by our research group, and an effective registration method was introduced in detail.

These papers [5–8] review a variety of image matching algorithms. These algorithms are mainly proposed for computer vision and medical image processing and are not suitable for matching different band images acquired by multispectral systems. Generally speaking, image registration methods can be placed into frequency domain methods and spatial domain methods [9]. The frequency domain methods, for example, the phase correlation method, are reliable with noise and intensity difference [10, 11]. From this perspective,

these frequency domain methods are suited for different spectral band images. However, the frequency method can only calculate the offset dislocation parameters between images. Because of imperfect mechanical integration, the cameras' optical axis is not strictly parallel. Only coarse-offset parameters can be obtained if we use the phase correlation method. The spatial domain algorithms can be further placed into area-based algorithms and feature-based algorithms [12]. The principle of the area-based method is to use the intensities of identical image windows to measure the similarity between two images. The following are well-known area-based methods: least squares matching (LSM) [13], maximum likelihood [14], statistical divergence [15], mutual information [16, 17], cross-correlation matching [18], and implicit similarity matching [19]. Although most of these area-based methods can obtain a high accuracy, they suffer from illumination difference [12] and are not suitable for different spectral band images.

Feature-based matching methods extract some feature points and match them using the information from pixels around feature points. Compared with area-based methods, most of feature-based methods remain invariant to illumination and viewpoint [20]. According to Mikolajczyk and Schmid's comparative study [21], the SIFT matching method [22] is one of the best feature-based methods. Compared with the images acquired by commercial cameras, the different spectral band images acquired by multispectral systems have a significant intensity difference, especially in vegetation area because its reflectance is sensitive to wavelength. The standard SIFT method compares every feature point in a reference image and all feature points in an input image to find its nearest neighbor and second nearest neighbor in the descriptor vector space, calculates the ratio of the distance from the nearest neighbor to the distance from the second nearest neighbor, determines whether or not to accept the feature point with the nearest descriptor according to the ratio, and then uses a registration model and random sample consensus (RANSAC) to eliminate false matches of initial matches. Due to the intensity difference, the correct matching feature point may not be the feature point with the nearest descriptor neighbor, and some descriptors of other feature points may be similar to the descriptor of the correct matching feature point. Therefore, in cases lacking man-made objects, the standard SIFT method may lead to two problems: one is that initial matches may be few, and the correct rate of initial matches may be low; the other is that their distribution may be uneven. However, the RANSAC algorithm is a learning technique to estimate parameters of a specific model by random sampling of observed data and uses a voting scheme to find the optimal fitting result; the voting scheme is based on an assumption that there are enough inliers (correct matches in this paper) to satisfy the specific model. When the inliers are few, the RANSAC usually performs badly [23]. Therefore, the standard SIFT method cannot be used for different spectral band images.

Some scientists improved the SIFT with different aspects [24–26] for remote sensing images. Gonçalves et al. [24]

proposed an automatic image registration method through image segmentation and SIFT. Hasan et al. [25] proposed remote sensing image registration via spatial relationship analysis on SIFT keypoints. Although these improved methods may obtain more correct matches, they must ensure that all the selected matches for spatial relationship analysis are correct. Based on SIFT and mutual information, Gong et al. [26] proposed a coarse-to-fine automatic image registration scheme for images with severe local deformation. These three improved methods are not suited for different band images acquired by the multispectral systems. Considering the nearly parallel integration of different cameras, this paper proposes an effective matching method for different band images. First we use one of the frequency methods, the phase correlation method, to calculate the coarse-offset relationship between one reference image and one input image. For any feature point in the reference image, we use this coarse-offset relationship to predict the location of its matching point in the input image. We then search the matching feature points near the predicted location other than all of the feature points of the input image.

This paper describes an airborne multispectral system, proposes an excellent registration method, and introduces this method in detail. The second section describes multispectral systems; the third section describes the methods; the fourth section introduces experiments; and the final section gives a conclusion.

## 2. The Multispectral Systems

In the recent years, many airborne multispectral systems, which consist of several identical monochrome cameras equipped with different bandpass filters, were developed by some scientists, such as Gorsevski and Gessler [2], Yang et al. [3], and Wu et al. [4], and their multispectral systems are, respectively, shown in Figures 1(a), 1(b), and 1(c). These multispectral systems have a common characteristic: their cameras are all nearly parallel to each other. Our research group also developed an airborne multispectral system, as shown in Figure 1(d), which is composed of a set of recorders, a ruggedized Getac B300 PC, four identical Hitachi KPF120CL monochrome cameras, and four bandpass filters. The four identical monochrome cameras, firmly installed and their optical axes parallel to each other, have the capability of obtaining 8-bit images with  $1392 \times 1040$  pixels and are, respectively, equipped with infrared (800 nm), red (650 nm), green (550 nm), and blue (450 nm) bandpass filters. As a result, this system has the flexibility to change filters for specific requirements. Because the four cameras are independent, they have the advantage that each camera can be individually adjusted for optimum focus and aperture setting. Due to imperfect mechanical integration, camera optical axes are not strictly parallel (the rotation vector  $R = [0, 0, 0] \pm [0.008, 0.008, 0.008]$ ) to each other, and it is impossible to align different band images optically or mechanically [27], so a registration method is needed.

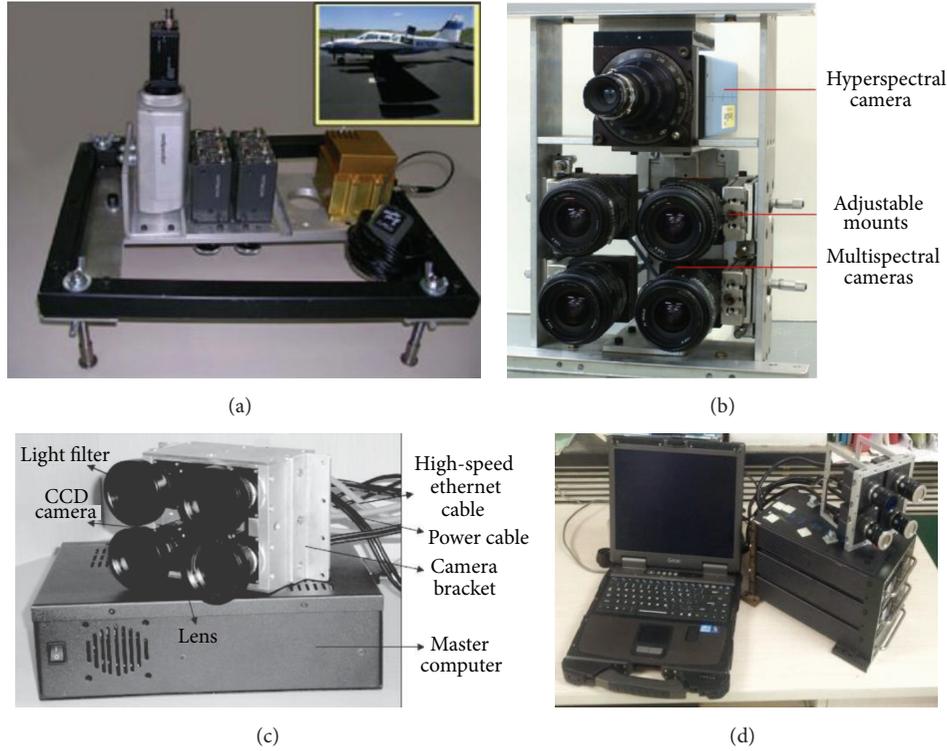


FIGURE 1: Multispectral systems.

### 3. Methods

**3.1. Phase Correlation.** In image processing, phase correlation is an image registration method and uses the fast frequency domain method to calculate the relative translative offset between two similar images. Given two images,  $f_1$  and  $f_2$ , that differ only by displacement  $(\Delta r, \Delta c)$ ,

$$f_2(x, y) = f_1(x - \Delta r, y - \Delta c). \quad (1)$$

After taking the Fourier Transform (FT) of both images, we get  $F_1$  and  $F_2$ . According to the translation property of the FT, we have the following relationship:

$$F_2(u, v) = F_1(u, v) \exp(-j2\pi(u\Delta r + v\Delta c)), \quad (2)$$

where  $u$  and  $v$  are the frequency variables in column and row. We can see that the shift in the spatial domain produces a phase difference in the frequency domain. Hence the normalized cross power spectrum of the images  $G$  is defined as

$$G(u, v) = \frac{F_1(u, v) F_2^*(u, v)}{|F_1(u, v) F_2^*(u, v)|} = \exp(-j2\pi(u\Delta r + v\Delta c)), \quad (3)$$

where  $F^*$  is the complex conjugate of  $F$ . By taking the Inverse Transform (IFT) of the normalized cross power spectrum,  $G$ , we can obtain the phase correlation, which is a two-dimensional delta function  $\delta(x - \Delta r, y - \Delta c)$  with a peak location corresponding to the displacement  $(\Delta r, \Delta c)$  between the two images.

During the past few years, many improved phase correlation methods, such as quadratic fitting method [28, 29], Gaussian fitting method [28, 29], Sinc fitting method [10], frequency domain masking method [11], and subspace identification extension method [30], have been proposed. Although these improved methods can achieve subpixel accuracy, they need a strict translation relationship between images. Due to the imperfect mechanical integration, the cameras' optical axes are not strictly parallel. There is no strict translation relationship between two images acquired by the multispectral systems. Therefore, these improved subpixel methods are meaningless for our multispectral system. We just need to use the standard phase correlation method to get a coarse-offset relationship for searching correct matches.

**3.2. SIFT.** The SIFT, one of the most powerful image matching methods, is proposed by Lowe [22]. It contains five major stages: scale-space extrema detection, keypoint location, orientation assignment, descriptor assignment, and keypoint matching.

**3.2.1. Scale-Space Extrema Detection.** In this stage, potential feature points are identified using a difference of Gaussian (DoG) function over all scales and image locations. For an image  $I(x, y)$ , the scale space is defined as  $L(x, y, \sigma)$ , which is constructed using a Gaussian Kernel  $G(x, y, \sigma)$  with different values of  $\sigma$ , as in

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y), \quad (4)$$

where  $*$  is the convolution operation in  $x$ ,  $y$ , and  $\sigma$  and  $G(x, y, \sigma)$  can be described as in

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}. \quad (5)$$

This operation is repeated using  $G(x, y, \sigma)$  to get a further blurred image  $L(x, y, k\sigma)$ . In order to detect stable feature points in scale space, DoG is calculated from two nearby blurred images separated by a constant multiplicative factor  $k$ , as in

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma). \end{aligned} \quad (6)$$

To find the extrema of  $D(x, y, \sigma)$ , each point is computed with its eight current scale neighbors, nine neighbors in the above scale, and nine neighbors in the below scale. The point is selected as a feature point only if it is larger or smaller than all of its neighbors.

**3.2.2. Keypoint Location.** To improve the stability, the feature points with low contrasts or poorly localized along the edges are rejected. The local detection can be refined using a fitting localization approach, proposed by Brown [6]. Local maximal can be detected via fitting the function  $D(x, y, \sigma)$  to nearby data.

**3.2.3. Orientation Assignment.** Each of selected feature points is assigned an orientation based on local image gradient directions. An orientation histogram which has 36 bins covering the  $360^\circ$  range of orientations is formed from the gradient orientations of sample points within a region around the feature point. The peak of the orientation histogram is assigned to the feature point. Therefore, the feature point is represented relative to the orientation and can achieve invariance to image rotation.

**3.2.4. Descriptor Assignment.** The feature point descriptor is a 128-dimensional vector which summarizes the magnitudes and orientations of the image gradient in a region around the feature point location. After being normalized to unity magnitude, it has a party invariance to image intensity.

**3.2.5. Keypoint Matching.** For each feature point in the reference image, its best candidate can be found by searching its nearest neighbor descriptor in all feature points of the input image. The nearest neighbor is the feature point whose descriptor has the minimum Euclidean distance from the given descriptor. Sometimes, for a feature point in a reference image, its correct matching point in the input image may not exist. In this case, the difference between the distance from its nearest neighbor and the distance from its second nearest neighbor is little. An effective method is to calculate the ratio of the distance from the nearest neighbor to the distance from the second nearest neighbor. If the ratio is smaller than a threshold, we accept it as the matching point of the feature point in reference image. Although this method

can reject many false matches, some correct matches may be removed. This situation becomes worse because of the nonlinear intensity difference between two different band images acquired by our multispectral system. Sometimes we cannot even acquire enough correct matches to calculate the transformation model by using the RANSAC.

**3.3. Motivation of Using Phase Correlation and SIFT.** Phase correlation is an image registration algorithm which uses a fast frequency domain approach to calculate the translation offset between two images. It has advantages in terms of its high computational efficiency, strong response to edges and salient pictures, and immunity to variation of image intensity compared with special domain algorithm and other frequency domain algorithms. It is suitable for registration across different spectral band images acquired by our multispectral system. Before we chose this method, we also tried many other methods such as the statistical divergence method, maximum likelihood method, mutual information method, cross-correlation method, the method by Keren et al. [31], and the frequency domain method by Vandewalle et al. [32] and found that only this method can get a correct and stable result.

The SIFT feature is a kind of local feature and maintains invariance of scaling and rotation and also keeps a certain degree of stability with change of brightness, the viewing change, affine transformation, and noise. According to a comparative study [21] on the performance assessment of different local descriptors, SIFT performs best for scale change, affine transformation, jpeg compression, rotation, image blur, and illumination change. We cannot find a better local descriptor method than SIFT.

Although the SIFT performs better than other contemporary descriptors, when we use it for the different band images acquired by a multispectral system directly, there are some problems: too many false matches exist in initial feature matches due to the significant difference in image intensity; only a few correct matches are obtained and their distribution is uneven. Applying the standard SIFT method directly cannot produce optimal results for the multispectral systems. Therefore, we need a reliable match algorithm to solve these problems.

**3.4. The Registration Method for Different Spectral Band Images.** Because the cameras of the multispectral systems are arranged in nearly parallel, we can assume that there is a coarse-offset relationship between different spectral band images. The coarse-offset relationship can be estimated using the phase correlation method mentioned in Section 3.1. Once the coarse-offset relationship is obtained, we can use a prediction strategy to greatly reduce search space. Assuming that the coarse-offset is  $(\Delta r, \Delta c)$ , for each SIFT feature point  $p_0$  with coordinate  $(x, y)$  in reference image  $I_0$ , its matching point  $p_i$  is near  $(x + \Delta r, y + \Delta c)$ . For searching the matching point,  $p_i$ , we only need to compare  $p_0$  and the feature points near  $(x + \Delta r, y + \Delta c)$  rather than the feature points all over the input image. Therefore, we can avoid eliminating some correct matches and producing some false matches through the standard SIFT method by using our method.

TABLE 1: The number of matches and correct rate.

	The standard method			Time	Our method			Time
	IM	CM	CR (%)		IM	CM	CR (%)	
Data 1	852	547	64.20	16.24 s	1243	1085	87.29	13.46 s
Data 2	208	44	21.15	16.06 s	614	523	85.18	13.50 s
Data 3	73	5	6.85	15.56 s	258	152	58.91	13.61 s
Data 4	80	7	8.75	15.98 s	196	127	64.80	13.74 s

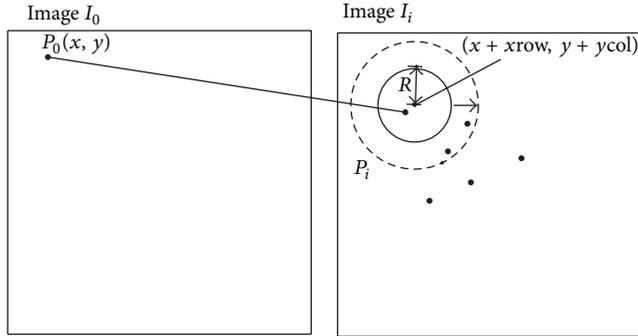


FIGURE 2: Schematic diagram of getting initial matches.

As shown in Figure 2, for any feature point  $p_0$  with coordinate  $(x, y)$  in reference image  $I_0$ , its matching point  $p_i$  is near  $(x + \Delta r, y + \Delta c)$  in input image  $I_i$ . We assume that the 128-dimensional descriptor vector of  $p_0$  is  $V_0$  and the feature points near  $(x + \Delta r, y + \Delta c)$  are  $p_{i0}, p_{i1}, \dots, p_{in}$  with descriptors  $V_{i0}, V_{i1}, \dots, V_{in}$ . We need to calculate the Euclidean distance between  $V_0$  and each one of  $V_{i0}, V_{i1}, \dots, V_{in}$  to find nearest and second nearest neighbor of  $V_0$ . If the ratio of the distance from the nearest neighbor to the distance from the second nearest neighbor is smaller than a given threshold (0.8, recommended by Lowe) and the distance between  $p_0$  and the feature point whose descriptor is the nearest one of  $V_{i0}, V_{i1}, \dots, V_{in}$  is less than the given initial radius  $R$  (10 pixels in this paper), we choose it as the matching point, or we think that there is no matching point of  $p_0$  in image  $I_i$ . The steps are as follows.

*Step 1.* Calculate the coarse-offset  $(\Delta r, \Delta c)$  between the reference image  $I_0$  and input image  $I_i$  using the phase correlation method.

*Step 2.* Give an initial radius to define the word “near.” In this paper, the initial radius,  $R$ , is 10 pixels, as the solid circle in Figure 2.

*Step 3.* Count the number of feature points near  $(x + \Delta r, y + \Delta c)$ ,  $N$ . If  $N \geq 2$ , continue; or enlarge the radius until  $N \geq 2$ . As shown in Figure 2, for only one feature point in the solid circle, we enlarge the radius; then for more than two feature points in the dotted circle, we stop enlarging the radius.

*Step 4.* Calculate the Euclidean distance between  $V_0$  and each one of  $V_{i0}, V_{i1}, \dots, V_{in}$  to find the nearest and second nearest neighbor of  $V_0$ .

*Step 5.* Calculate ratio of the distance from the nearest neighbor to the distance from the second nearest neighbor. Assume that the nearest neighbor of  $V_0$  is  $V_{ij}$ . If the ratio is lower than a given threshold (0.8, recommended by Lowe) and the distance between  $p_{ij}$  and  $p_0$  is less than the given initial radius  $R$  (10 pixels in this paper), we choose  $p_0$  and  $p_{ij}$  as an initial match.

After these steps, the number of obtained matches is larger and the correct rate is higher. There are still some false matches; the RANSAC can be used to eliminate them.

## 4. Experiment and Analysis

Our multispectral system was mounted on a metal protective box installed on an airship, named ASQ-HAA380, which is developed by our research group. On August 16, 2014, the research group carried a flight experiment in Haibei Tibetan Autonomous Prefecture, Qinghai province, China. The flying height is 300 meters, and the image size is  $1392 \times 1040$ . The experimental scenario is shown in Figure 3.

*4.1. Registration Results between Visible Bands and Infrared Bands.* We select four image pairs acquired by our multispectral system, as shown in Figure 4. Because it is not difficult to register images among the three kinds of visible bands, we just show the results that every image pair includes a kind of visible band image (red band image) and an infrared band image. There are many artificial objects in Data 1, shown in Figures 4(A) and 4(a), fewer artificial objects in Data 2, shown in Figures 4(B) and 4(b), and no artificial objects in Data 3, shown in Figures 4(C) and 4(c), and Data 4, shown in Figures 4(D) and 4(d). Figures 4(A), 4(B), 4(C), and 4(D) show the registration results by the standard SIFT method; Figures 4(a), 4(b), 4(c), and 4(d) show the registration results using our method. Their statistical results as shown in Table 1 are 852, 208, 73, and 80 initial matches (IM) and 547, 44, 5, and 7 correct matches (CM) in Data 1, Data 2, Data 3, and Data 4, respectively, using the standard SIFT method; there are 1243, 614, 258, and 196 IM and 1085, 523, 152, and 127 CM in Data 1, Data 2, Data 3, and Data 4, respectively, by our method. The phenomena illuminate that the SIFT matches decrease with the decrease of artificial objects regardless of IM or CM, no matter which method is used. In comparison with Figures 4(A) and 4(B), Figures 4(a) and 4(b) have more IM and CM. In Figure 4(B), the matches distribute only in the artificial road area, and it also illustrates that matches increase with the increase of artificial objects in image pairs.



FIGURE 3: Experimental scene and equipment installation.

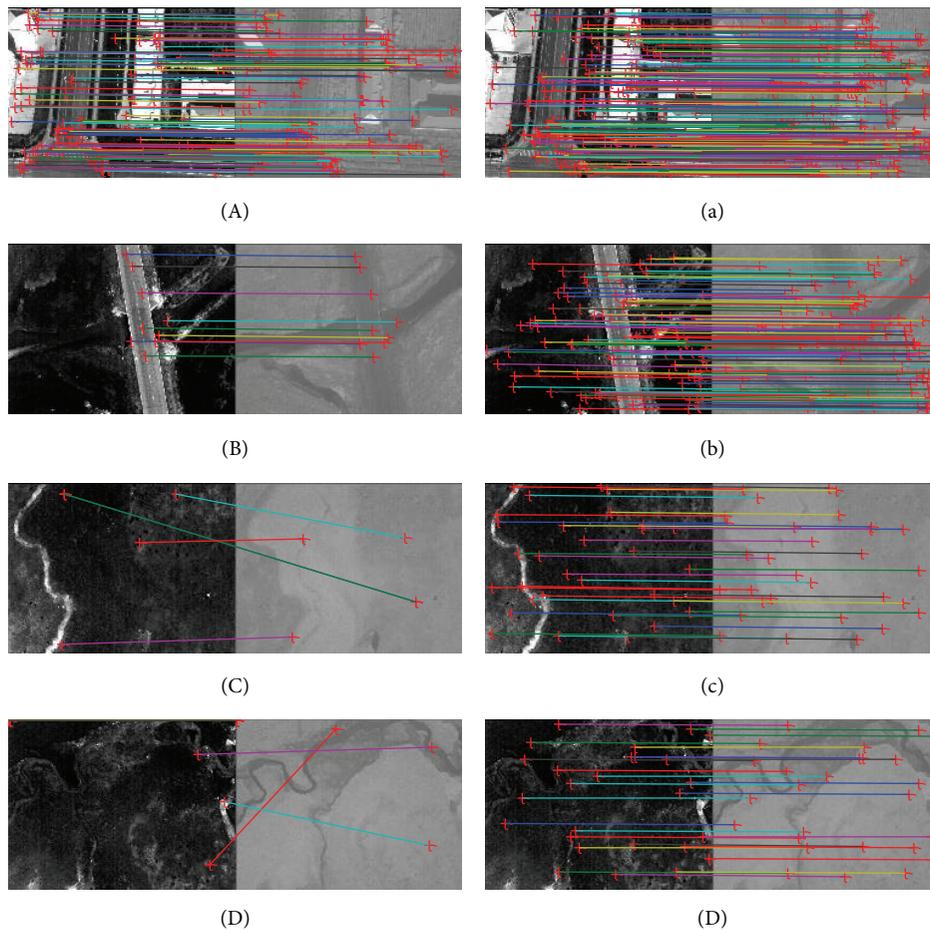


FIGURE 4: Image registration between one kind of visible band image (red band images, the left of every pair) and the infrared band image (the right of every pair): (A), (B), (C), and (D) registered using the standard SIFT method; (a), (b), (c), and (d) registered by our method. To show these images clearly, only one-fifth of matches are shown.

However, the matches distribute nearly all over the image pair in Figure 4(b). It can be concluded that more CM are obtained, and they distribute better using our method. The correct rate (CR) of initial matches in Data 1, Data 2, Data 3, and Data 4 is, respectively, 64.20%, 21.15%, 6.85%, and 8.75% using the standard SIFT method and 87.29%, 85.18%, 58.91%, and 64.80% by our method. It can be concluded that the CR also decrease with the decrease of artificial objects, and the CR of our method is obviously higher than the CR of

the standard SIFT method, especially in Data 2, Data 3, and Data 4. Because Data 3 and Data 4 contained no artificial objects, the CM is less, and the CR is low when using the standard SIFT method; the RANSAC cannot be used directly. So, the standard SIFT method cannot get a correct result, as shown in Figures 4(C) and 4(D). However, there still are many CM obtained in Data 3 and Data 4 using our method, and their CR is still high. So, the correct registration results can also be obtained by our method, and the CM distribute

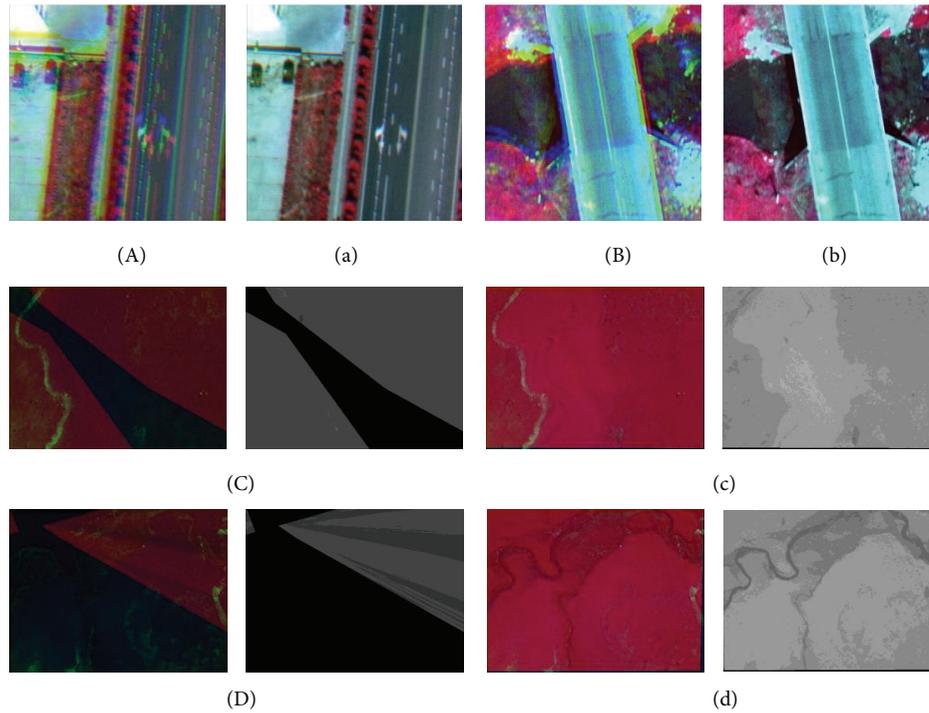


FIGURE 5: The CIR composite of multispectral images and its single infrared band: (A) and (a) enlarged partial regions of unregistered and registered composite of Data 1; (B) and (b) enlarged partial regions of unregistered and registered composite of Data 2; (C)-left and (D)-left, the registered composite of Data 3 and Data 4 using the standard SIFT method; (C)-right and (D)-right, the single infrared band of Data 3 and Data 4 using the standard SIFT method; (c)-left and (d)-left, the registered composite of Data 3 and Data 4 using our method; (c)-right and (d)-right, the single infrared band of Data 3 and Data 4 using our method.

well. All these experiments are tested in a ThinkPad L440, with a 2.5 GHz Intel i5 CPU, 4G memory. As shown in Table 1, the time consumption of the standard SIFT method is about 16 s. And the time consumption of our method is about 13.5 s, about 2.5 s less than the standard SIFT method.

#### 4.2. Registration Results for Multispectral Image Composition.

The process of multispectral image composition is as follows: three empty output matrixes are created at the same pixel size as the reference image (a red band image). For any one of the three, using homography model [27], the coordinates of each pixel in the output image matrix are transformed to determine their corresponding location in the original input image. However, because of small differences in the CCD sensors and orientations of the cameras in the aluminum frame, this transformed cell location will not directly overlay a pixel in the input matrix. Interpolation is used to assign a Digital Number (DN) value to the output matrix pixel determined on the basis of the pixel values that surround its transformed position in the input image matrix. Some common interpolation methods can be chosen, including nearest neighbor algorithm, bilinear interpolation, cubic convolution, and inverse distance weighting (IDW). The nearest neighbor algorithm is adopted because it assigns the closest pixel value in the input image to the new pixel and does not change the original data values. Of course, we may use other complex interpolation methods if necessary.

Figures 5(A) and 5(B) show enlarged partial regions of CIR composite (infrared, red, and green) of unregistered Data 1 and Data 2, respectively; Figures 5(a) and 5(b) show enlarged partial regions of CIR composite of registered Data 1 and Data 2, respectively. There are severe dislocations between different bands of the unregistered multispectral image. In contrast, these dislocations are missing in the registered multispectral image. Because of a lack of artificial objects in these images in Data 3 and Data 4, there is a great difference between visible SIFT features and infrared SIFT features, and the CR of initial matching point pairs is low using the standard SIFT method. Therefore, it cannot get a correct registration infrared band in the multispectral image, as shown in Figures 5(C)-right and 5(D)-right, and correct CIR composites, as shown in Figures 5(C)-left and 5(D)-left. However, our method can get an excellent result with many CM and a high CR. So, the single infrared bands, as shown in Figures 5(c)-right and 5(d)-right, and the CIR composites, as shown in Figures 5(c)-left and 5(d)-left, are correct.

In order to have a quantitative measure of registration error, the red bands of the composition multispectral images are taken as inference bands, and the registered infrared bands of the composition multispectral image are taken as input bands; our method, or the standard SIFT method, is used to get correct matches. For all matches, the coordinates of the feature points in the reference band image are, respectively,  $(x_{r1}, y_{r1}), (x_{r2}, y_{r2}), \dots, (x_{rn}, y_{rn})$  where the coordinates of the corresponding points in the input band

TABLE 2: The errors between the visible bands and infrared bands of a multispectral image.

	The standard method			Our method		
	$R_x$	$R_y$	$R_t$	$R_x$	$R_y$	$R_t$
Data 1	1.275	1.446	1.361	1.273	1.445	1.359
Data 2	1.692	1.572	1.632	1.623	1.495	1.559
Data 3	/	/	/	2.387	3.117	2.752
Data 4	/	/	/	2.618	2.862	2.740

image are, respectively,  $(x_{i1}, y_{i1}), (x_{i2}, y_{i2}), \dots, (x_{in}, y_{in})$ , and  $r$  and  $i$  represent the points in reference bands and in input bands, respectively. With  $n$  as the number of the point pairs,  $j$  as the serial number, the root mean square error in  $x$  direction, the root mean square error in  $y$  direction, and the total root mean square error, ( $R_x$ ,  $R_y$ , and  $R_t$ ) can be calculated as the following formulas:

$$\begin{aligned}
 R_x &= \sqrt{\frac{1}{n} \sum_{j=1}^n (x_r - x_i)^2}, \\
 R_y &= \sqrt{\frac{1}{n} \sum_{j=1}^n (y_r - y_i)^2}, \\
 R_t &= \sqrt{R_x^2 + R_y^2}.
 \end{aligned} \tag{7}$$

These errors indicate how good the registration is between the reference band image and the input band image. The smaller these errors are, the higher the quality of the multispectral data is. Table 2 indicates that our method and the standard SIFT method have nearly the same quality in Data 1; in Data 2, the error of our method is 0.073 pixels lower than the standard SIFT method; in Data 3 and Data 4, the standard SIFT method cannot get correct results, but our method can. Its total error is about 2.7 pixels, which is still very low.

## 5. Conclusion

We developed an airborne multispectral system using four changeable bandpass filters, four identical monochrome cameras, a set of recorders, and a ruggedized PC. For this multispectral system and other systems similar to it, we proposed an effective registration method. In the case of no artificial objects, the difference in intensity between visible band images and infrared band images acquired by our system or other multispectral systems is significant, and the standard SIFT registration approach cannot solve the matching problem. Thus, the correct multispectral images cannot be obtained. To solve this problem, this paper proposed an effective method based on the common structural characteristic of the multispectral systems, phase correlation, and the SIFT. It uses the phase correlation method to estimate the coarse-offset parameters, uses the SIFT to acquire feature points, uses the coarse-offset parameters to predict the coordinate of its matching point in an input image for any feature point in a reference image, and then compares the feature points near

the predicted coordinate in the input image with the feature point in the reference image for getting a matching point. Compared with the standard method, our method not only improves the registration performance but also solves the matching problem between an infrared image and a visible image in cases lacking man-made objects.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work is supported by program for National Natural Science Foundation of China (NSFC 41071255), the National Key Technologies R&D Program (2012BAH31B01), and the Key Project B of Beijing Natural Science Foundation (KZ201310028035). This work is funded by Specialized Research Fund for the Doctoral Program of Higher Education (SRFDP20131108110005) and the Importation and Development of High-Caliber Talents Project of Beijing Municipal Institutions (CIT&TCD20150323).

## References

- [1] N. Oppelt and W. Mauser, "Airborne visible/infrared imaging spectrometer AVIS: design, characterization and calibration," *Sensors*, vol. 7, no. 9, pp. 1934–1953, 2007.
- [2] P. V. Gorsevski and P. E. Gessler, "The design and the development of a hyperspectral and multispectral airborne mapping system," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 64, no. 2, pp. 184–192, 2009.
- [3] C. Yang, J. H. Everitt, M. R. Davis, and C. Mao, "A CCD camera-based hyperspectral imaging system for stationary and airborne applications," *Geocarto International*, vol. 18, no. 2, pp. 71–80, 2003.
- [4] J. Wu, Y. S. Cao, H. S. Yang, and Z. G. Liu, "Automatic registration of high-resolution multispectral imageries released by low-altitude unmanned airship," *Journal of Remote Sensing*, vol. 16, no. 3, pp. 625–643, 2012.
- [5] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.
- [6] L. G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, 1992.
- [7] L. M. G. Fonseca and B. S. Manjunath, "Registration techniques for multisensor remotely sensed imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 9, pp. 1049–1056, 1996.

- [8] A. Goshtasby, *2-D and 3-D Image Registration for Medical, Remote Sensing and Industrial Applications*, Wiley-Interscience, New York, NY, USA, 2005.
- [9] J. Li, Q.-M. Peng, and Z.-H. Fan, "A survey of sub-pixel image registration methods," *Journal of Image and Graphics*, vol. 13, no. 11, pp. 2070–2075, 2008.
- [10] H. Foroosh, J. B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 11, no. 3, pp. 188–200, 2002.
- [11] 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 Transactions on Geoscience and Remote Sensing*, vol. 39, no. 10, pp. 2235–2243, 2001.
- [12] A. Sedaghat, M. Mokhtarzade, and H. Ebadi, "Uniform robust scale-invariant feature matching for optical remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4516–4527, 2011.
- [13] A. Gruen, "Adaptive least squares correlation: a powerful image matching technique," *South African Journal of Photogrammetry, Remote Sensing and Cartography*, vol. 14, no. 3, pp. 175–187, 1985.
- [14] W. Li and H. Leung, "A maximum likelihood approach for image registration using control point and intensity," *IEEE Transactions on Image Processing*, vol. 13, no. 8, pp. 1115–1127, 2004.
- [15] Y. He, A. B. Hamza, and H. Krim, "A generalized divergence measure for robust image registration," *IEEE Transactions on Signal Processing*, vol. 51, no. 5, pp. 1211–1220, 2003.
- [16] H.-M. Chen, M. K. Arora, and P. K. Varshney, "Mutual information-based image registration for remote sensing data," *International Journal of Remote Sensing*, vol. 24, no. 18, pp. 3701–3706, 2003.
- [17] J. P. Kern and M. S. Pattichis, "Robust multispectral image registration using mutual-information models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 5, pp. 1494–1505, 2007.
- [18] A. Goshtasby, S. H. Gage, and J. F. Bartholic, "A two-stage cross correlation approach to template matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 3, pp. 374–378, 1984.
- [19] Y. Keller and A. Averbuch, "Multisensor image registration via implicit similarity," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 794–801, 2006.
- [20] L. Juan and O. Gwun, "A comparison of SIFT, PCA-SIFT and SURE," *International Journal of Image Processing*, vol. 3, no. 4, pp. 143–152, 2009.
- [21] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.
- [22] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [23] A. Hast, J. Nysjö, and A. Marchetti, "Optimal RANSAC—towards a repeatable algorithm for finding the optimal set," *Journal of WSCG*, vol. 21, no. 1, pp. 21–30, 2013.
- [24] H. Gonçalves, L. Corte-Real, and J. A. Goncalves, "Automatic image registration through image segmentation and SIFT," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 7, pp. 2589–2600, 2011.
- [25] M. Hasan, X. Jia, A. Robles-Kelly, J. Zhou, and M. R. Pickering, "Multi-spectral remote sensing image registration via spatial relationship analysis on sift keypoints," in *Proceedings of the 30th IEEE International Geoscience and Remote Sensing Symposium (IGARSS '10)*, pp. 1011–1014, IEEE, July 2010.
- [26] M. Gong, S. Zhao, L. Jiao, D. Tian, and S. Wang, "A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 7, pp. 4328–4338, 2014.
- [27] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, UK, 2nd edition, 2003.
- [28] G. A. Thomas, "Television motion measurement for DATV and other applications," NASA STI/Recon Technical Report N 88: 13496, 1987.
- [29] I. E. Abdou, "Practical approach to the registration of multiple frames of video images," in *Visual Communications and Image Processing '99*, vol. 3653 of *Proceedings of SPIE*, pp. 371–382, International Society for Optics and Photonics, San Jose, Calif, USA, January 1998.
- [30] W. S. Hoge, "A subspace identification extension to the phase correlation method [MRI application]," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 277–280, 2003.
- [31] D. Keren, S. Peleg, and R. Brada, "Image sequence enhancement using sub-pixel displacement," in *Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 742–746, June 1988.
- [32] P. Vandewalle, S. Süsstrunk, and M. Vetterll, "A frequency domain approach to registration of aliased images with application to super-resolution," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 071459, pp. 1–14, 2006.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

