

Journal of
Applied Remote Sensing

**Elastic registration for airborne
multispectral line scanners**

Feng Li
ChuanRong Li
LingLi Tang
Yi Guo

Elastic registration for airborne multispectral line scanners

Feng Li,^{a,*} ChuanRong Li,^a LingLi Tang,^a and Yi Guo^b

^aAcademy of Opto-Electronics, Chinese Academy of Sciences, Key Laboratory of Quantitative Remote Sensing Information Technology, Beijing 100094, China

^bCSIRO Computational Informatics, New South Wales 1670, Australia

Abstract. The multispectral line scanner is one of the most popular payloads for aerial remote sensing (RS) applications. Scanners with large field of view (FOV) improve efficiency in Earth observation. Small-volume instruments with a short focal length and a large FOV, however, may bring a problem: different nonlinear warping and local transformation exist between bands. Alignment accuracy of bands is a criterion impacting product quality in RS. A band-to-band elastic image registration method is proposed for solving the problem. Rather than ignoring the intensity variation and carrying out an intensity-based registration between bands straightforwardly, we construct feature images and use them to conduct an intensity-based elastic image registration. In this method, the idea of the inverse compositional algorithm is employed and expanded when dealing with local warping, and a smoothness constraint is also added in this procedure. Experimental results show that the proposed band-to-band registration method works well both visually and quantitatively. The outstanding performance of the method also encourages potential applications for other new types of airborne multispectral imagers. © *The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: [10.1117/1.JRS.8.083614](https://doi.org/10.1117/1.JRS.8.083614)]

Keywords: band-to-band registration; affine transform; multispectral scanners; unmanned aerial vehicle.

Paper 13549P received Jan. 2, 2014; revised manuscript received May 4, 2014; accepted for publication May 16, 2014; published online Jun. 13, 2014.

1 Introduction

Remote sensing (RS) is developing toward high spatial resolution, high time resolution, and high spectral resolution. Multispectral images of line scanners are the main type of data acquired by RS sensors in Earth observation. Unlike space mission applications, platforms of aerial applications suffer more effects such as pitch, yaw, and roll. Boresight error consists of yaw error, pitch error, and roll error. For an airborne line scanner, i.e., pushbroom scanner, pitch error will make the whole line data see either forward or backward instead of nadir. Roll error will shift the whole line data either to the left or the right. Yaw error may cause the whole line data to rotate along the flight direction. If the boresight error is constant during the flight, it is easy to correct. However, in reality the boresight error varies from time to time caused by air turbulence, wind, and platform vibration. These irregular effects can bring a problem for airborne multispectral line scanners: raw RS multispectral bands may be not innately registered. Alignment accuracy of multispectral bands from airborne scanners is a criterion impacting product quality in RS. Many scholars have been investigating to solve the band-to-band registration problem.

There are dozens of intensity-based image registration methods, but none of them can be used to solve the multispectral image registration problem in a straightforward manner. The reason is that different spectral characteristics of ground lead to intensity variation between bands. Especially, visible light images and infrared light images of the same scene will be regarded as totally different images in term of intensity. There are some algorithms¹⁻³ to solve the

*Address all correspondence to: Feng Li, E-mail: feng_li@aliyun.com

band-to-band registration problem for the Enhanced Thematic Mapper Plus and moderate-resolution imaging spectroradiometer sensors. In general, the scale-invariant feature transform (SIFT) algorithm⁴ gives excellent matching results for visible image pairs. For infrared and visible image pairs, however, the number of mismatched SIFT keypoints increases rapidly, and this decreases the SIFT-based registration accuracy drastically. Therefore, authors in Ref. 5 proposed a new scale restriction criteria for keypoint matching with better performance. Moreover, there are also some algorithms based on mutual information⁶ and statistical learning techniques⁷ to solve the band-to-band rigid/global registration problem. A correlation and Hough transform-based method of automatic image registration was proposed in Ref. 8. The authors claimed that the method can globally register a pair of images with different spectral contents. Common elastic image registration methods^{9,10} are based on the intensity of the images. Since intensity of the same region varies according to different spectral characteristics of ground, those elastic methods often fail. Moreover, high computational cost of Periaswamy's^{9,10} method limits its application for RS images. Therefore, a new and efficient elastic band-to-band registration is needed to solve the problem.

Recently, a new type of multispectral line scanner developed by Changchun Institute of Optics in China is composed of four off-axis aspheric reflectors in order to achieve a wide field of view (FOV) with a short focal length. The new multispectral scanner can achieve an FOV of 61.93 and IFOV of 0.18 mrad with four bands: blue (420 to 520 nm), green (520 to 600 nm), red (630 to 690 nm), and near-infrared (760 to 900 nm). A picture of the multispectral line scanner is shown in Fig. 1. The physical dimensions of the scanner are about $150 \times 220 \times 200$ (mm). This kind of scanner is vulnerable to the problem of different nonlinear warping and inconsistent local transformation between bands. This is caused by air turbulence, self-propelled vibration of airborne platforms, and the systematic design of the multispectral line scanner. As mentioned earlier, common global image registration methods are not suitable for this problem. One example image taken by this scanner is shown in Fig. 2. We can see that each band suffers different nonlinear warping, although the flight was carried out under good weather conditions. Based on our survey results, most of the band-to-band registration methods can only solve the misalignment globally, and they fail to register bands with local warps shown in Fig. 2. This is problem is not brought only by this particular multispectral line scanner. Another new type of spectrometer was proposed and the technique is abbreviated as IRIS.^{11,12} Based on the IRIS design scheme, all multispectral bands are projected on a single conventional detector array, and this kind of spectrometer may also suffer the above problem when it is used for airborne applications.

In this paper, a new elastic band-to-band registration method is proposed to solve the above problem. The rest of this paper is organized as follows. Feature images are introduced in Sec. 1, then a new elastic image registration method is proposed in Sec. 2. In Sec. 4, we demonstrate the

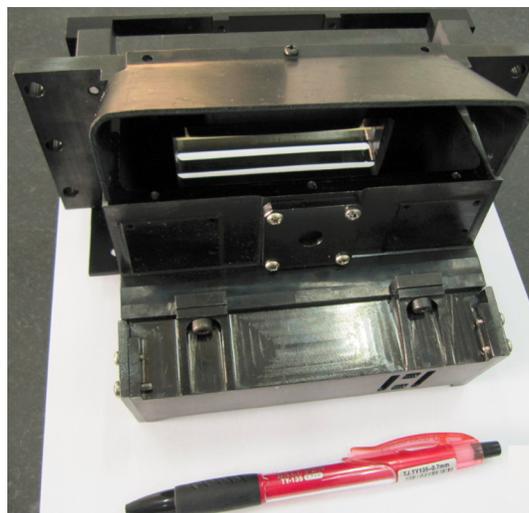


Fig. 1 Picture of the multispectral line scanner.

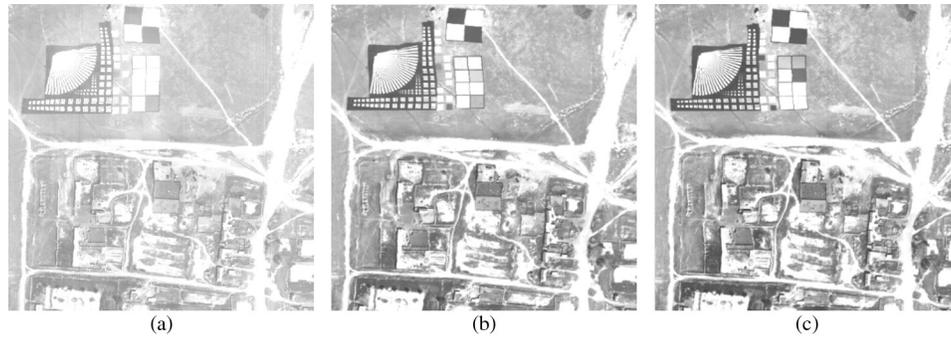


Fig. 2 An example of captured bands with the new multispectral line scanner. (a) Blue band, (b) green band, and (c) red band.

high accuracy of registration between bands with the new method for some real flight data. Finally, conclusions are provided in Sec. 5.

2 Methodology

2.1 Feature Images

Nonrigid warps between multispectral bands require elastic registration methods considering both light intensity variation and local warping. Human eyes are very sensitive to edges rather than smooth areas of a scene. This is the reason humans can easily recognize the same region, although the region of intensity is different across different wavelength ranges. Considering the characteristic of human eyes, we mimic its edge detection by introducing a gradient operator to explore the similarities within multispectral images. We define feature images as follows:

$$F^k(x, y) = \sqrt{\left(\frac{\partial I^k(x, y)}{\partial x}\right)^2 + \left(\frac{\partial I^k(x, y)}{\partial y}\right)^2}, \quad (1)$$

where $I^k(x, y)$ denotes the k 'th channel of multispectral images. After normalizing the feature image for each band, edges of images will be emphasized. The feature images of the bands in Fig. 2 are shown in Fig. 3. Rather than carry out registration between bands straightforwardly, we use feature images constructed for each band to conduct an intensity-based elastic image registration. Before the registration step, a histogram-matching step is applied to the reference feature image in order to match the histogram of the target feature image in terms of intensity.

2.2 Feature Images-Based Elastic Image Registration

Because the feature images summarize the major transition of the regions by exploring the similarity within bands, we align the corresponding feature images instead of multispectral images



Fig. 3 An example of feature images of the bands shown in Fig. 2.

with a new elastic image registration method. We denote the reference and source feature images as $T(x, y)$ and $S(u, v)$, respectively.

In this paper, the idea of the inverse compositional algorithm^{13,14} is adopted and expanded in dealing with both rigid and local warping with high efficiency. To improve the efficiency, we warp the target image toward the source image so that a Hessian matrix can be obtained outside of the iterations, which avoids the clumsy step-by-step warping with a Taylor series expansion as in Ref. 9.

2.2.1 Global affine model

We use a seven-dimensional vector to describe the warp and the brightness variations between the source and reference images. The motion between them can be modeled by an affine transform

$$T(x, y) = S(u, v) + m_7 = S(m_1x + m_2y + m_5, m_3x + m_4y + m_6) + m_7, \quad (2)$$

where m_1, m_2, m_3 , and m_4 are the linear affine parameters, m_5 and m_6 are the translation parameters, and m_7 denotes the variation of intensity between the two images; $u = m_1x + m_2y + m_5$ and $v = m_3x + m_4y + m_6$.

Equation (2) can also be denoted as $T(x, y) = S[W(x, y; M)] + m_7$, where

$$W(x, y; M) = \begin{bmatrix} m_1 & m_2 & m_5 \\ m_3 & m_4 & m_6 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3)$$

is used to express the affine warp to the coordinate frame and M is the affine parameters matrix

$$\begin{bmatrix} m_1 & m_2 & m_5 \\ m_3 & m_4 & m_6 \\ 0 & 0 & 1 \end{bmatrix}.$$

In general, the process of image registration is to keep the source image aligned with the reference image; therefore, warping the source image toward the reference image step-by-step in each iteration is very common. In contrast, in our method, the reference image is warped toward the source image step-by-step and the inverse of the final warping parameters is applied to the source image to complete the registration.

We set $\vec{m} = [\tilde{m}_1, \tilde{m}_2, \tilde{m}_3, \tilde{m}_4, \tilde{m}_5, \Delta m_6, m_7]^T$ and

$$\Delta M = \begin{bmatrix} \tilde{m}_1 & \tilde{m}_2 & \tilde{m}_5 \\ \tilde{m}_3 & \tilde{m}_4 & \tilde{m}_6 \\ 0 & 0 & 1 \end{bmatrix},$$

where $\Delta M = M^{-1}$. Therefore, by warping the reference image $T(u, v)$ with the parameter set ΔM , the reference image will be aligned with the source image $S(x, y)$. Then, we have

$$T[W(u, v; \Delta M)] = T(\tilde{m}_1u + \tilde{m}_2v + \tilde{m}_5, \tilde{m}_3u + \tilde{m}_4v + \tilde{m}_6) = S(u, v) + m_7. \quad (4)$$

The cost function is as follows:

$$E(\vec{m}) = \sum_{u, v \in \Omega} [T(\tilde{m}_1u + \tilde{m}_2v + \tilde{m}_5, \tilde{m}_3u + \tilde{m}_4v + \tilde{m}_6) - S(u, v) - m_7]^2, \quad (5)$$

where Ω denotes the image region.

The goal is to minimize the above cost function to calculate ΔM , which is the inverse of the affine parameter matrix M . If $T[W(u, v; \Delta M)]$ is expanded using a first-order truncated Taylor series, then

$$E(\vec{m}) \approx \sum_{u,v \in \Omega} [T + (\tilde{m}_1 u + \tilde{m}_2 v + \tilde{m}_5 - u)T_u + (\tilde{m}_3 u + \tilde{m}_4 v + \tilde{m}_6 - v)T_v - S - m_7]^2. \quad (6)$$

T and S are symbolized as lexicographically ordered vectors of $T(x, y)$ and $S(x, y)$ of length N^2 , where N^2 denotes the size of images. T_u and T_v are the spatial derivatives of $T(u, v)$ and are symbolized as lexicographically ordered vectors of length N^2 . Then the error function can be simplified as follows:

$$E(\vec{m}) = \|k - \vec{c}^T \vec{m}\|_2^2, \quad (7)$$

where k and vector \vec{c} are given by

$$k = S - T + uT_u + vT_v \quad (8)$$

and

$$\vec{c} = (uT_u, vT_u, uT_v, vT_v, T_u, T_v, -e)^T, \quad (9)$$

where $-e$ denotes a vector in which all components are -1 . The size of \vec{c} is $7 \times N^2$.

Then the error function in Eq. (7) can be minimized by differentiating $E(\vec{m})$ with respect to the unknowns \vec{m} as

$$\frac{dE(\vec{\Delta m})}{d\vec{\Delta m}} = -2\vec{c}[k - \vec{c}^T \vec{\Delta m}]. \quad (10)$$

Then

$$\vec{\Delta m} = [\vec{c}\vec{c}^T]^{-1}[\vec{c}k]. \quad (11)$$

Note that $\vec{\Delta m}$ in Eq. (11) is the warping parameter set for $T(x, y)$ to be aligned with $S(u, v)$, so the affine transform matrix M for $S(u, v)$ is obtained as follows:

$$M = \Delta M^{-1}. \quad (12)$$

The initial M is a 3×3 identity matrix. Initially, m_7 is set to 0 and is updated at each iteration. After several iterations, the source image will be warped by M to match the reference image.

Note that, compared with Periaswamy's method, $[\vec{c}\vec{c}^T]$ (the Hessian matrix of the reference image) is a constant matrix, because there is nothing in this matrix that depends on iteratively warped S , so it can be precomputed. The additional step is to calculate Eq. (12) at each iteration. Compared with the calculation of $[\vec{c}\vec{c}^T]$, the computational cost of the additional step, i.e., $M = \Delta M^{-1}$ is very small, because ΔM is a 3×3 matrix only. N^2 is set to be the number of pixels in the reference image, and then the saved cost in each iteration is $O(7^2 N^2)$. In this way, the efficiency of Periaswamy's algorithm has been improved in processing the global affine warp.

In dealing with large distortion between the source image and the target image, a coarse to fine scheme is adopted. A Gaussian pyramid as shown in Fig. 4 is built for both the source image and the reference image, and the global affine parameters are calculated from the coarsest level.

2.2.2 Local shift model

Two images can be coarsely aligned after the global affine transform mentioned above. Now, a local warping between local regions needs to be considered. As in many other methods, the two images are divided into many boxes to deal with the local warping individually. The smaller the box size, the better the complex nonrigid warping can be approximated by translation. Therefore, a three-dimensional vector is adopted to describe the translation and brightness variation between the source image box and the reference image box.

Let $BT(x, y)$ and $BS(u, v)$ denote the reference and source image boxes, respectively. The motion between them can be modeled by a shift transform

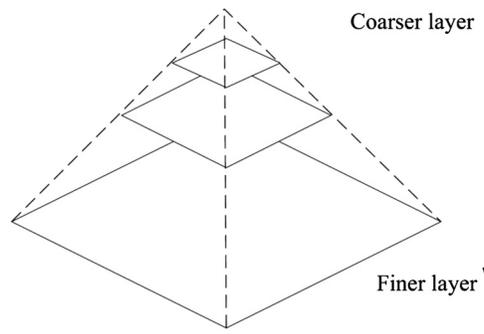


Fig. 4 The registration starts from coarser layers to finer layers in the image pyramid.

$$BT(x, y) = BS(x + m_5, y + m_6) + m_7 = BS(u, v) + m_7, \quad (13)$$

where $u = x + m_5$ and $v = y + m_6$. Then, the above equation can be rewritten as

$$BT(u - m_5, v - m_6) = BS(u, v) + m_7. \quad (14)$$

We can either shift the source box a particular amount to align it with the reference box, or shift the reference box in the inverse direction by the same amount to keep them aligned. The latter strategy is adopted to improve the efficiency as described in Sec. 2.2.1. Then, the cost function is defined by

$$E_{\text{box}}(\vec{m}) = \sum_{u, v \in \Omega} [BT(u - m_5, v - m_6) - BS(u, v) - m_7]^2, \quad (15)$$

where Ω denotes the box and $\vec{m} = [m_5, m_6, m_7]^T$.

The goal is to minimize this expression to calculate the translation parameters and the intensity variation. $BT(u - m_5, v - m_6)$ is expanded using a first-order truncated Taylor series, thus

$$E_{\text{box}}(\vec{m}) \approx [BT - m_5 BT_u - m_6 BT_v - BS - m_7]^2, \quad (16)$$

where BT_u and BT_v are the spatial derivatives of $BT(u, v)$ in lexicographically ordered vectors, BT and BS are symbolized as lexicographically ordered vectors of $BT(x, y)$ and $BS(x, y)$, respectively. The error function can be simplified as follows:

$$E_{\text{box}}(\vec{m}) = \|k - \vec{c}^T \vec{m}\|_2^2, \quad (17)$$

where k and \vec{c} are given by

$$k = BT - BS \quad (18)$$

and

$$\vec{c} = [BT_u, BT_v, e]^T, \quad (19)$$

where e denotes a vector in which all components are 1.

The cost function in Eq. (17) can be minimized by differentiating $E_{\text{box}}(\vec{m})$ with respect to the unknowns \vec{m} by

$$\frac{dE_{\text{box}}(\vec{m})}{d\vec{m}} = \sum_{u, v \in \Omega} -2\vec{c}[k - \vec{c}^T \vec{m}]. \quad (20)$$

Then

$$\vec{m} = [\vec{c}\vec{c}^T]^{-1}[\vec{c}k]. \quad (21)$$

As we can see, the Hessian matrix $[\vec{c}\vec{c}^T]^{-1}$ is a constant matrix, because there is nothing in this matrix that depends on the shifted source box, so it can be precomputed. Once it is

calculated, it can be saved in memory for use in each iteration. On the contrary, the vector $[\vec{c}k]$ has to be calculated in each iteration. After a few iterations, the best \vec{m} is obtained for Eq. (17) by making $E_{\text{box}}(\vec{m})$ minimum between two boxes.

2.2.3 Smoothness constraint

Usually, the motion parameters vary smoothly in most parts of the image. We assume that neighbor pixel locations may have constant/smooth motion parameters. This provides a very useful constraint for obtaining meaningful local alignment parameters. During every local registration iteration introduced above, to avoid a large difference between neighbor parameter sets and to express the movement parameters with intensity variation only, a smoothness constraint is applied to these parameters.

As in Refs. 9 and 15, a similar smoothness constraint is set up. The error function of every pixel is described as follows:

$$E_{\text{pix}}(\vec{m}) = E_{xy}(\vec{m}) + E_{\text{sm}}(\vec{m}), \quad (22)$$

where $E_{xy}(\vec{m})$ is defined as in Eq. (17) centralizing (x, y) as

$$E_{xy}(\vec{m}) = [k - \vec{c}^T \vec{m}]^2 \quad (23)$$

with k , \vec{c} , and \vec{m} the same expression as shown in Eq. (17). The new term $E_{\text{sm}}(\vec{m})$ denotes the smoothness constraint as in⁹

$$E_{\text{sm}}(\vec{m}) = \sum_{e=5}^7 \zeta_e \left[\left(\frac{\partial m_e}{\partial x} \right)^2 + \left(\frac{\partial m_e}{\partial y} \right)^2 \right], \quad (24)$$

where ζ_e is a prespecified positive constant which adjusts the relative weight given to the smoothness constraint on the parameter m_e . The larger ζ_e , the weaker the rigidity, i.e., more weight on the smoothness. By differentiating Eq. (22) with respect to the model parameters and setting the gradient to zero, we can minimize the error function in Eq. (22) as follows:

$$\frac{dE_{\text{pix}}(\vec{m})}{d\vec{m}} = \frac{dE_{xy}(\vec{m})}{d\vec{m}} + \frac{dE_{\text{sm}}(\vec{m})}{d\vec{m}} = 0. \quad (25)$$

The derivative of $E_{xy}(\vec{m})$ is

$$\frac{dE_{xy}(\vec{m})}{d\vec{m}} = -2\vec{c}[k - \vec{c}^T \vec{m}]. \quad (26)$$

To compute the derivative of $[dE_{\text{sm}}(\vec{m})]/(d\vec{m})$, the method used in Refs. 9 and 15 is chosen by using the horizontal and vertical derivatives of $m_e(x, y)$ as

$$\frac{\partial m_e(x, y)}{\partial x} = m_e(x, y) - m_e(x + 1, y), \quad \frac{\partial m_e(x, y)}{\partial y} = m_e(x, y) - m_e(x, y + 1), \quad (27)$$

where $m_e(x, y)$ is the e 'th image registration parameter of the coordinate of (x, y) .

Then, Eq. (24) can be written as

$$E_{\text{sm}}[m_e(x, y)] = \zeta_e \{ [m_e(x, y) - m_e(x + 1, y)]^2 + [m_e(x, y) - m_e(x, y + 1)]^2 \}. \quad (28)$$

Therefore, $\{dE_{\text{sm}}[m_e(x, y)]\}/[dm_e(x, y)]$ can be approximated as

$$\begin{aligned} \frac{dE_{\text{sm}}[m_e(x, y)]}{dm_e(x, y)} &= 2\zeta_e \{ m_e(x, y) - m_e(x + 1, y) + m_e(x, y) - m_e(x, y + 1) \} \\ &= 4\zeta_e [m_e(x, y) - \bar{m}_e(x, y)], \end{aligned} \quad (29)$$

where $\bar{m}_e(x, y) = [m_e(x + 1, y) + m_e(x, y + 1)]/2$, or the local average of m_e at the coordinate of (x, y) .

Using vector notation, the derivative of $E_{sm}(\vec{m})$ is denoted by discrete approximations⁹

$$\frac{dE_{sm}(\vec{m})}{d\vec{m}} = 2L(\vec{m} - \bar{\vec{m}}), \quad (30)$$

where $\bar{\vec{m}}$ is the component-wise average of \vec{m} over a small spatial neighborhood and L is a 3×3 diagonal matrix with diagonal elements ζ_e , and zeros off diagonal. L is the parameter matrix that weights the error in the local shift model registration relative to the departure from smoothness. ζ_e should be small if the parameter set calculation from E_{xy} is accurate and large if there are errors with parameter set calculation caused by the intensity variation setting $\{[dE_{xy}(\vec{m})]/d\vec{m}\} + \{[dE_{sm}(\vec{m})]/d\vec{m}\} = 0$, we have

$$\vec{m}^{(j+1)} = (\vec{c}\vec{c}^T + L)^{-1}(\vec{c}k + L\bar{\vec{m}}^{(j)}). \quad (31)$$

At each iteration j , $\bar{\vec{m}}^{(j)}$ can be estimated from the current $\vec{m}^{(j)}$. The initial $\vec{m}^{(0)}$ can be estimated from Eq. (21).

Here, $\zeta_e = 1$ in our experimental tests, and the selection of the size of local box Ω depends on particular images. The smaller the box size, the more the elastic warping is fixed; however, the weaker the constraint that is applied. A larger box size means that more pixels provide constraints on the mean square error minimization. Therefore, it shows more rigid characteristics. A smaller box size brings more freedom with less constraint, but it also brings more error because the matrix $[\vec{c}\vec{m}^T]$ may not be invertible. There is a natural tradeoff in choosing the box size.

3 Overview and Implementation Details

Here, the overall structure of the new elastic image registration for multispectral images will be reviewed. First, feature images of the multispectral bands are created. One of the feature images is selected as a reference image and other feature images as source images. Then, the image pyramid is constructed from the two input feature (source and reference) images. The new image registration algorithm starts from the coarsest layer images. The global affine image registration is applied to these two coarsest source and reference feature images. After a certain number of global iterations, a parameter set ΔM will be calculated. Because $\Delta M = M^{-1}$, M can be calculated, which is used to warp the source image to the reference image for this current layer. Then it brings this warped source image and the reference image into the local shift registration procedure. In each local iteration, two images are divided into boxes, and the shift parameters are calculated for each coordinate position.

After each local shift registration iteration, a number of smoothness constraint iterations follow to avoid some unexpected errors. Once the local shift registration procedure and the smoothing procedure finish, the warping parameters within that layer are interpolated by a factor of 2 for the next finer-layer image until reaching the resolution of the original source and reference images. In order to avoid the artifacts of the bilinear interpolation in each warping procedure for the source image pyramid, an accumulated warp is finally applied to the original source image pyramid layer at each level. For example, if 15 global iterations were to be adopted, there would be 15 steps in warping the source image layer toward the reference image layer at each scale. Therefore, 15 times bilinear interpolation would be sequentially applied to the same source image layer. Then, the blur artifacts would be extremely severe. Therefore, rather than continuously warping and interpolating the same source image layer, the warping parameter set is accumulated step-by-step and the overall warping parameter set is applied to the original source image layer in a single warp at the end for each source layer image.

Based on the feature images of both the reference and the source, the pixel translation of each position: m_5 and m_6 can be calculated between the reference feature image and the source feature image. Finally, we warp (i.e., a grid step plus a bilinear interpolation step) the source band toward the reference band with the precalculated translation values m_5 and m_6 . The above operation should be routinely carried out for each band. Eventually, the whole multispectral band-to-band registration is completed.

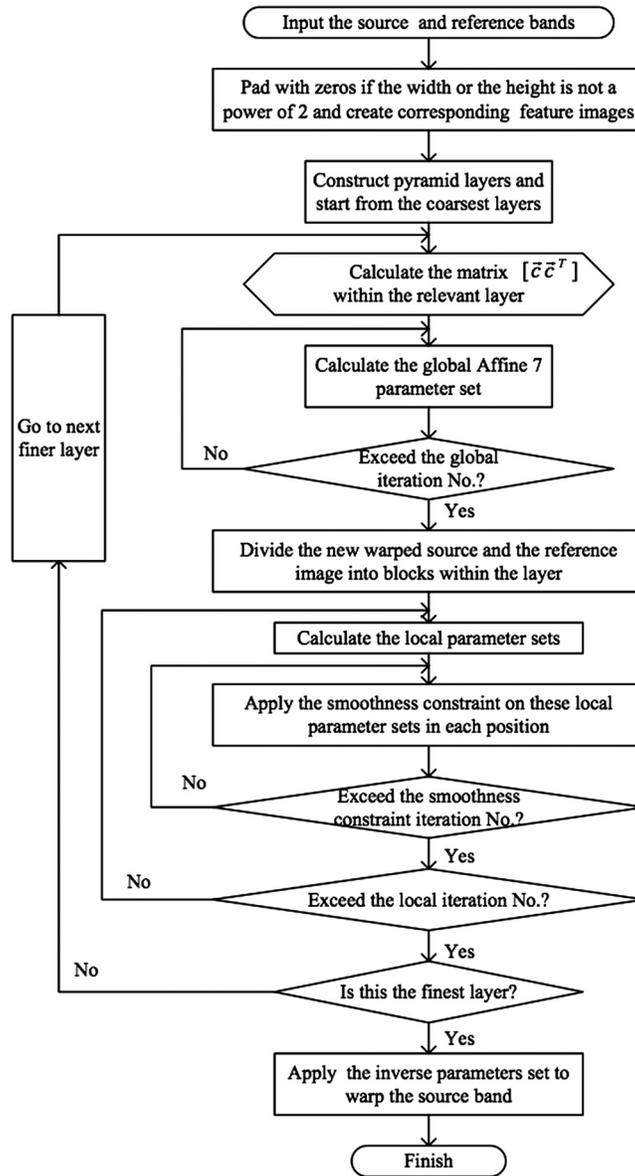


Fig. 5 Flow chart of our proposed registration method.

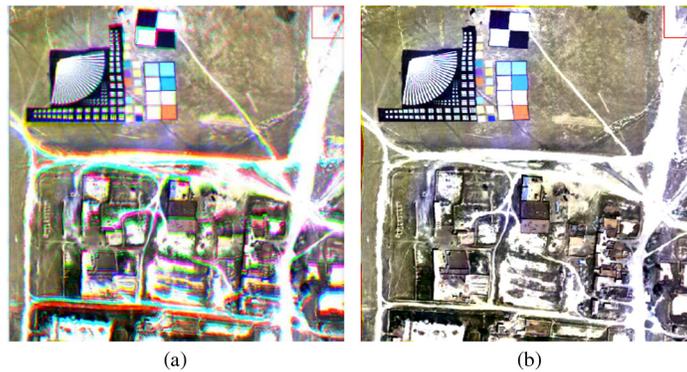


Fig. 6 The true color images before and after the elastic registration with the test data shown in Fig. 2. (a) Before registration and (b) after registration.

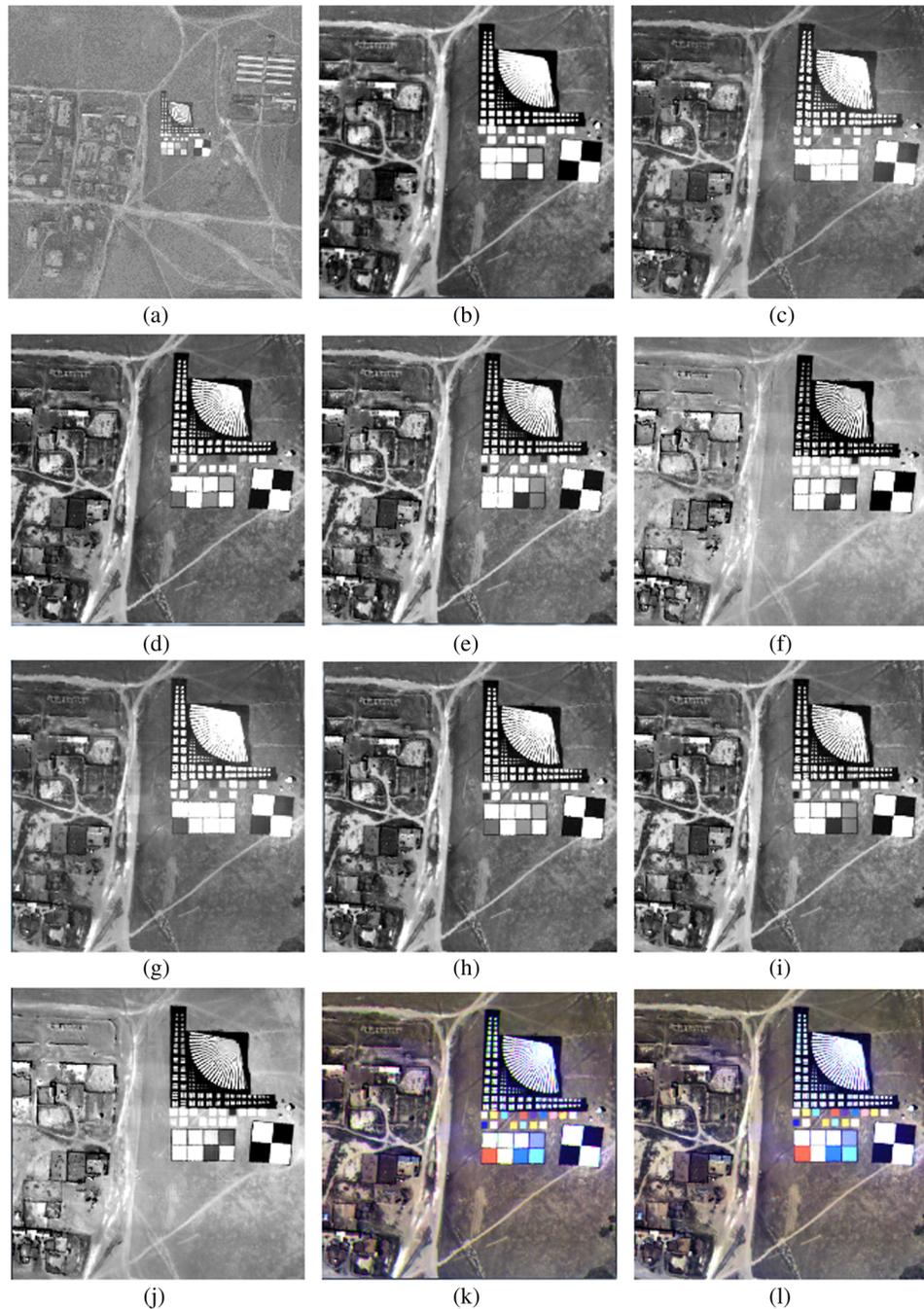


Fig. 7 An example by registering an image captured by an array camera onboard. (a) Whole image from array camera, (b) subregion of (a), (c) blue band, (d) green band, (e) red band, (f) near-infrared band, (g) registered blue band, (h) registered green band, (i) registered red band, (j) registered near-infrared band, (k) true color image (before), and (l) after registration.

To help the readers better understand the whole process of our method, a flow chart is shown in Fig. 5. Moreover, MATLAB code and other details of our algorithm can be found at <http://code.google.com/p/b2bregistration>.

4 Experimental Results

In order to evaluate the performance of the proposed method, we tested our method with real flight data. By reforming an unmanned aerial vehicle (UAV), multipayloads such as the new type

Table 1 Numerical comparison results before and after the registration.

Bands	Horizontal shift average value in pixels		Vertical shift average value in pixels	
	Before registration	After registration	Before registration	After registration
Blue band	4.79	0.20	0.71	0.09
Green band	3.55	0.11	1.02	0.08
Red band	3.32	0.09	0.66	0.07
Near-infrared band	4.03	0.21	0.59	0.25

of wide-view multispectral line scanner introduced above, a hyperspectral camera and a panchromatic array camera are integrated into the vehicle. The flight was conducted in a testing site (for calibration and validation) near BaoTou city of China on September 3, 2011. The total flight time was over 10 h with three different flight altitudes. One of the test data sets acquired by the new multispectral line scanner without the geometric correction procedure is shown in Fig. 2. The true color image shown in Fig. 6(a) indicates the misalignment between bands. By setting the red band image Fig. 2(c) as a reference, we register the blue band Fig. 2(a) and the green band Fig. 2(b) to the reference, respectively. The true color image after the elastic registration is shown in Fig. 6(b). We can see that the true color image shows no misalignment between bands. Note that the true color image still suffers from some geometric distortion. This is because geometric distortion exists in the red band, which is the reference.

To decrease the geometric distortion in the reference image, we took images captured by the array camera in another flight track as the reference. Moreover, a geometric correction step was used to achieve the goal. In our flight test, payloads were equipped on an airborne photoelectric stabilized platform. Given the position and orientation system (POS) parameter and the compensation parameters from the airborne photoelectric stabilized platform, a geometric correction step can be applied to the raw image data to decrease the geometric error brought by the UAV attitude. Images captured by the panchromatic array camera generally suffer less local warps caused by the vibration of the platform than those from line scanner cameras. The reason is that each line data may suffer different POS and photoelectric stabilized platform parameters during the exposure period. A further test image captured by the array camera on the same day but in another track of the flight is shown in Fig. 7(a) with size of 1024×1024 pixels. An enlarged region of (a) is shown in Fig. 7(b). Images of the same region of the multispectral data after geometric correction are shown in Figs. 7(c)–7(f) of blue, green, red, and near-infrared, respectively. From Figs. 7(c)–7(f), we can see that there are still some nonlinear local warps between bands, although they are much less than those in the data shown in Fig. 2. By setting the image from the array camera as the reference band, the same region of the registered images of blue, green, red, and near-infrared are shown in Figs. 7(g)–7(j), respectively. Moreover, to help the readers have a visual comparison, the true color image of the same region before and after the elastic registration procedure are shown in Figs. 7(k) and 7(l), respectively. We can see that the true color image of Fig. 7(l) contains less misalignment between registered bands and less geometric distortions. In this experiment, $\Omega = 40 \times 40$ was selected as the box size for solving local warps.

To demonstrate the performance of our band-to-band registration method quantitatively, we carried out a numerical comparison by evaluating the misalignment between the registered bands and the panchromatic image before and after registrations. We randomly selected 11 testing points on edges or corners in the whole images to ensure that these points are identifiable for all bands. This is because testing points from smooth regions cannot be uniquely identified by either human eyes or evaluation methods. After selecting the testing points, a paraboloid surface fit method¹⁶ was adopted to evaluate the registration accuracy in subpixels. The mean of the horizontal and vertical misalignments of the 11 points between bands and the reference (the panchromatic image captured by the onboard array camera) is shown in Table 1. Both visual and numerical comparison show that our elastic band-to-band registration method works well for solving different nonlinear warping and local translations between bands.

5 Conclusions

In this paper, a new elastic band-to-band registration method is proposed for solving local warps between bands. Gradient-based feature images are constructed as intermediate products to assist the registration. Instead of registering multispectral bands straightforwardly, we apply a new intensity-based elastic registration method to register those feature images to calculate their corresponding warping parameters. Experimental results of real data have shown that the proposed method can correct local warps in band-to-band registration with satisfactory performance. With a growing number of real applications of the new multispectral line scanner of large FOV and short focal length, this band-to-band registration method will certainly attract more attention.

Acknowledgments

Feng Li acknowledges the support of NSFC under Grant 41371415. Yi Guo is supported by the CSIRO Minerals Down Under Research Flagship and Capability Development Fund.

References

1. J. Barker and J. Seiferth, "Landsat thematic mapper band-to-band registration," in *Int. Geoscience and Remote Sensing Symposium, 1996 (IGARSS'96) Remote Sensing for a Sustainable Future*, Lincoln, Nebraska, Vol. 3, pp. 1600–1602, IEEE (1996).
2. K. Yang et al., "Modis band-to-band registration," in *Proc. IEEE 2000 Int. Geoscience and Remote Sensing Symposium, 2000 (IGARSS 2000)*, Honolulu, Hawaii, Vol. 2, pp. 887–889, IEEE (2000).
3. R. Eastman, J. Le Moigne, and N. Netanyahu, "Research issues in image registration for remote sensing," in *IEEE Conf. on Computer Vision and Pattern Recognition, 2007 (CVPR'07)*, Minneapolis, Minnesota, pp. 1–8, IEEE (2007).
4. D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision* **60**(2), 91–110 (2004).
5. Z. Yi, C. Zhiguo, and X. Yang, "Multi-spectral remote image registration based on sift," *Electron. Lett.* **44**(2), 107–108 (2008).
6. J. Kern and M. Pattichis, "Robust multispectral image registration using mutual-information models," *IEEE Trans. Geosci. Remote Sens.* **45**(5), 1494–1505 (2007).
7. T. Kim and M. Choi, "Fast multi-spectral image registration based on a statistical learning technique," *Proc. SPIE* **7810**, 78100C (2010).
8. H. Gonçalves et al., "CHAIR: automatic image registration based on correlation and Hough transform," *Int. J. Remote Sensing* **33**(24), 7936–7968 (2012).
9. S. Periaswamy and H. Farid, "Elastic registration in the presence of intensity variations," *IEEE Trans. Med. Imaging* **22**(7), 865–874 (2003).
10. S. Periaswamy and H. Farid, "Elastic registration with partial data," in *Biomedical Image Registration*, pp. 102–111, Springer, Berlin, Heidelberg (2003).
11. A. Harvey et al., "Spectral imaging in a snapshot," *Proc. SPIE* **5694**, 110–119 (2005).
12. A. Gorman, D. Fletcher-Holmes, and A. Harvey, "Generalization of the Lyot filter and its application to snapshot spectral imaging," *Opt. Express* **18**(6), 5602–5608 (2010).
13. S. Baker and I. Matthews, "Lucas-Kanade 20 years on: a unifying framework," *Int. J. Comput. Vision* **56**(3), 221–255 (2004).
14. G. Ye et al., "Efficient multi-image registration with illumination and lens distortion correction," in *IEEE Int. Conf. on Image Processing, 2005 (ICIP 2005)*, Genoa, Italy, Vol. 3, pp. III-1108–III-1111, IEEE (2005).
15. B. Horn, *Robot Vision*, MIT Press, Cambridge, MA (1986).
16. S. S. Gleason, M. A. Hunt, and W. B. Jatko, "Subpixel measurement of image features based on paraboloid surface fit," *Proc. SPIE* **1386**, 135–144 (1991).

Feng Li received his PhD degree from the University of New South Wales in 2009. He joined CSIRO in the same year and developed algorithms for the applications of compressive sensing to radio astronomy. Currently, he is working at the Academy of Opto-electronics, Chinese

Academy of Sciences. His research interests are in general image processing, image restoration, and compressive sensing.

ChuanRong Li is a professor at the Academy of Opto-electronics, Chinese Academy of Sciences. His research is mainly on remote sensing data processing, geometric correction, image processing, and remote sensing data calibration and validation.

LingLi Tang is a professor at the Academy of Opto-electronics, Chinese Academy of Sciences. Her research is mainly on remote sensing data preprocessing, SAR data processing, image processing, and remote sensing data calibration and validation.

Yi Guo obtained a BE degree in electrical engineering from North China University of Technology in 1998 and a PhD degree from University of New England in 2008. He has been with the Commonwealth Science and Industry Organisation (CSIRO) as a research scientist since 2008. His research focuses on machine learning and computational statistics for big data.