A COARSE-TO-FINE BAND REGISTRATION FRAMEWORK FOR MULTI/HYPERSPECTRAL REMOTE SENSING IMAGES CONSIDERING CLOUD INFLUENCE

Xuhui Zhao[1], Zhi Gao[1],[*] Wenbo Sun[2], Fei Wen[1]

[1] School of Remote Sensing and Information Engineering, Wuhan University, 430079 Wuhan, China -
(zhaoxuhui, wenfei)@whu.edu.cn, gaozhinus@gmail.com
[2] School of Electronic Information, Wuhan University, 430079 Wuhan, China - wenbosun@whu.edu.cn

Commission III, WG III/4

ABSTRACT:

Band registration is one of the most critical steps in the production of multi/hyperspectral images and determines the accuracy of applications directly. Because of the characteristics of imaging devices in some multi/hyperspectral satellites, there may be a time difference between bands during push-broom imaging, which leads to displacements of moving clouds with respect to the ground. And a large number of feature points may gather around cloud contours due to the high contrast and rich texture, resulting in building a transformation more suitable for moving clouds and making ground objects ghosted and blurred. This brings a big challenge for registration methods based on feature extraction and matching. In this paper, we propose a novel coarse-to-fine band registration framework for multi/hyperspectral images containing moving clouds. In the coarse registration stage, a cloud mask is generated by grayscale stretching, morphology and other operations. Based on this mask, a coarse matching of cloud-free regions is performed to eliminate large misalignment between bands. In the refinement stage, low-rank analysis and RASL (Robust Alignment by Sparse and Low-rank decomposition) are used to optimize the rank of coarse results to achieve fine registration between bands. After experiments on a total of 102 images (83 cloudy images and 19 cloud-free images with all 32 bands) from Zhuhai-1 hyperspectral satellite, our method can achieve a registration accuracy of 0.6 pixels in cloudy images, 0.41 pixels in cloud-free images, which is enough for subsequent applications.

1. INTRODUCTION

With the continuous development of science and technology, more multi/hyperspectral sensors have been equipped on earth observation satellites, which effectively expands the ability to observe the earth through rich spectral information. However, due to the characteristics of imaging devices in some multi/hyperspectral satellites, the imaging time of the same ground object may vary in different bands during push-broom imaging, which leads to misalignment between bands. Therefore, band registration is one of the most important steps in the production of multi/hyperspectral data and there are mainly two types of registration methods at present. One is the method that takes into account satellite imaging models (hereinafter referred to as model-based method), and the other is the method that uses computer vision (hereinafter referred to as vision-based method). Model-based methods use attitude, orbit, and time data of satellites during imaging to build a geometric position-ing model between ground points and image pixels. The relative relationship between bands can be restored using this point-to-point model, which may achieve a high registration accuracy. But it needs many satellite operational parameters that are not easy to get and use. And the calculation for the satellite ima-ging model is complicated and tedious. Vision-based methods usually estimate a transformation by extracting and matching tie points[1] in different bands and uses it for alignment. Com-pared with model-based methods, it is more straightforward
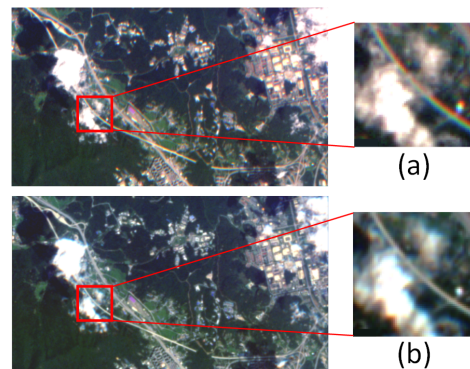

Figure 1. Influence of moving clouds on multi/hyperspectral band registration in the cloudy image. (a) Registration results of band red, green and blue using the vision-based method. Roads become rainbow-like due to the movement of clouds. (b) Results after eliminating cloud motion with our methods.

and requires fewer parameters and computations. However, as mentioned before, there may be a slight movement of mov-ing clouds in different bands with respect to the ground due to the imaging time difference. And many tie points may gather around cloud contours because of the strong contrast and rich texture when extracting and matching features. Transforma-tions based on these tie points are inaccurate and will lead to a poor registration or even failure for ground objects while a good accuracy for clouds as shown in Fig.1. It brings a big challenge

---

for vision-based methods, but many do not take into account this impact of moving clouds on band registration. Aiming at the problem of poor registration for ground objects in cloudy multi/hyperspectral images, we propose a novel coarse-to-fine band registration framework that eliminates the influence of moving clouds.

The main structure of this paper is as follows: Some work related to this topic is introduced in Section 2. Then the coarse-to-fine band registration framework is described in Section 3 in detail. Extensive experiments are performed on data from Zhuhai-1 hyperspectral satellite and results are in Section 4. Conclusions are summarized in Section 5.

2. RELATED WORK

2.1 Cloud Detection in Remote Sensing Images

Clouds cover about 50% of the earth's surface (Randall, Corsetti, 1989), (King et al., 2013), which indicates that clouds inevitably appear in most remote sensing images. Cloud detection is a subject that has been studied for years, many methods have been proposed for general or specific scenarios. In general, it can be divided into two categories: methods based on image grayscale (hereinafter referred to as grayscale-based method) and methods based on physical spectrum (hereinafter referred to as spectrum-based method).Grayscale-based methods use computer vision methods for cloud detection based on the grayscale value of pixels rather than the physical spectral reflection characteristics, which are quite simple and easy to use. It can be further divided into learning-based methods and non-learning methods. Different convolutional neural networks (Ozkan et al., 2018),(Li et al., 2019a),(Shao et al., 2019) were used for cloud detection and got some good results. As for non-learning methods, it involves broader technologies. A nonnegative matrix factorization method was used for cloud detection(Li et al., 2019b). (Xu et al., 2019) studied noise-adjusted principal components transform for removing clouds in optical remote sensing images.(Wang et al., 2016),(Deng et al., 2016) tried to propose a general and automatic method for detecting clouds. Besides using a single image,(Wen et al., 2018),(Zhang et al., 2019) used low-rank analysis to remove clouds in remote sensing image sequences and proposed a coarse-to-fine framework for detecting and inpainting clouds from other time serial images. Spectrum-based methods are mainly used in multispectral images. It can be divided into three categories: threshold, radial transfer, and statistical. Among them, the threshold method is a commonly used method because of its simplicity and small calculation amount (Hagolle et al., 2010),(Jedlovec et al., 2008),(Zhu et al., 2015). Clouds are considered cold water vapor which has higher reflectivity and lower temperature than the ground. Simple visible and infrared window thresholding methods perform quite well in cloud detection(Ackerman et al., 1998). It may have higher detection accuracy in complex scenes while a narrower application because of its requirements on the data type. And it may be hard to distinguish objects with similar spectral characteristics such as snow from clouds.

2.2 Multi/Hyperspectral Image Band Registration

Currently, two kinds of methods are widely used for multi/hyperspectral band registration : model-based methods and vision-based methods. Model-based methods usually build a rigid geometric imaging model for simulation based on satellite operational parameters during the observation process and solve the geographic location of ground objects pixel by pixel. It may achieve high accuracy in some complex scenarios, while it has to use many operational parameters that are not easy to get and may change often. On the other hand, the process of model-based methods is quite sophisticated, which hinders the versatility for different sensors and platforms, such as small drones.(Jiang et al., 2013) achieved high accuracy band-to-band registration for ZY-3 satellite by virtual re-imaging. (Jiang et al., 2019) proposed a band-to-band registration method based on positioning consistency constraint for Zhuhai-1 hyperspectral satellite. A pixel in a certain band can be calculated to the corresponding ground point by establishing the imaging model, and then the ground point can be calculated back to other bands. Theoretically, points in the other bands should coincide with the points in the reference band, but they will not actually coincide completely. This difference between bands can be used to find the relationship for band registration. Vision-based methods usually extract corresponding feature pairs (such as points, lines, etc.) and use these relationships to fit a transformation (such as affine, homography, etc.) between bands. (Pan et al., 2011) used a Harris corner detector (Harris et al., 1988) for feature detection and matching. And they also used OpenMP technology to implement the parallel band-to-band registration. (Yi et al., 2008) modified the SIFT(Scale-invariant Feature Transform) (Lowe, 1999) for better performance on multispectral images, which they called SR-SIFT. (Ma et al., 2016) also modified the SIFT for remote sensing image registration. Besides modifying existing features in computer vision, (Ye, Shan, 2014) proposed a two-step registration method for multispectral images. They used SR-SIFT for primarily matching. And then a Harris corner detector and an LSS(Local Self-Similarity) descriptor were used for fine matching. It achieves good results in cloud-free images, but cannot eliminate the cloud movement in bands. There are still some other methods that do not leverage on feature extraction and matching. (Dong et al., 2011) provided a registration method based on optical flow, it is suitable for small images but needs lots of calculations when images are large. (Wang et al., 2018) proposed a deep learning framework for remote sensing image registration. However, it just works on optical images and needs a lot of labeled data to train the network. (Peng et al., 2012) proposed a new method for alignment of batch images, which is called RASL (Robust Alignment by Sparse and Low-rank decomposition). They regard image alignment as an optimization problem of the rank of a matrix and achieve some great results. Inspired by this work, (Hu et al., 2014) used rank minimization for multiangle multi/hyperspectral images. But the initialization of optimization is complex which makes it hard for practical use. Generally compared with model-based methods, vison-based methods are more universal and can be used for alignment when we do not know parameters for building an imaging model.

As a summary, we propose a coarse-to-fine multi/hyperspectral band registration framework for cloudy images to eliminate the influence of moving clouds on registration. We adopt grayscale-based methods for cloud detection because of its simplicity and good performance, which can be used for panchromatic images. And we achieve coarse alignment by vision-based methods for the high efficiency and wide applicability. Finally, we use RASL for fine registration based on coarse stage results.

3. THE PROPOSED FRAMEWORK

Fig.2 shows the main flow of the coarse-to-fine framework for multi/hyperspectral band registration.In the coarse stage, two

processings are performed on the input band simultaneously, one is to generate the cloud mask for feature filtering based on thresholding and morphological operations, the other is to enhance the contrast of cloud-free part for a better feature extraction. After extraction and filtering, the filtered features are matched to find a coarse band-wise transformation. In the refinement stage, results from the coarse registration are used as initial inputs to RASL for optimization, and finally a more accurate registration result is obtained.
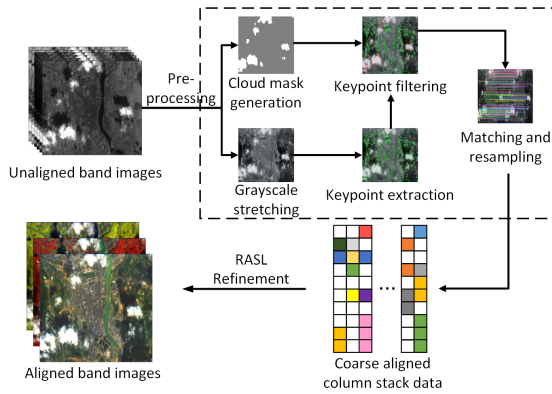

Figure 2. Flowchart of the proposed coarse-to-fine registration framework for multi/hyperspectral images.

3.1 Coarse Band Registration

In cloudy images, a large amount of cloud-contour feature points are included because of the strong contrast and rich texture around cloud contours. And due to the brightness of clouds, the ground becomes darker after grayscale stretching, which further aggravates this situation and builds a model more suitable for clouds, resulting in a poor registration accuracy for ground objects. In the coarse stage, our goal is to increase the proportion of ground feature points in total as much as possible, so as to find the model that fits most matches (ground features) for alignment. We generate a cloud mask to filter cloud-contour points and perform grayscale stretching to enhance the contrast of the ground for more feature points. In practice, the cloud mask and filtered points may not be absolutely accurate due to errors, but it is enough for the initialization of the refinement stage. For finding the cloud mask, the input image is pre-processed to make the bright part brighter and the dark part darker. A new stretching function is proposed as Eq.1:

$$y = \begin{cases} 2^{m-1}x^m & 0 \le x \le 0.5, \\ (-2)^{m-1} + 1 & 0.5 < x \le 1 \end{cases} \quad (1)$$

Where x and y are the input and output normalized grayscale value, m is a coefficient used to adjust the stretching effect. The larger m indicates more obvious effects. We set $m = 4$ for experiments in this paper. After obtaining the stretched image, OTSU binarization (Otsu, 1979) is performed on it. Then a dilation operation is used to remove fragmented pixels and noise, and multiple erosion operations are used to expand the cloud range. For the ability to describe details we use small morphological kernels. And to filter "fake" feature points around cloud contours as possible, a "greedy strategy" is adopted for a complete coverage of the cloud. On the other hand, due to limited sensor quantization levels, the grayscale contrast of remote sensing images is usually low, which may lead

to failure and mismatching of algorithms based on grayscale information. Therefore, grayscale stretching is usually performed on images before feature extraction. But with the existence of clouds, grayscale value and contrast in dark cloud-free parts(compared with clouds) reduces significantly after ordinary grayscale stretching, which affects the quantity and quality of features. For a better feature extraction, another effective grayscale stretching algorithm is designed by combining logarithmic stretching and linear stretching. The core grayscale transformation is as shown in Eq.2.

$$g(i,j) = d \cdot \log_{1+k}(1 + \frac{ck}{d} + \frac{(d-c)k}{(b-a)d}(f(i,j)-a)) \quad (2)$$

Where $f(i,j)$ and $g(i,j)$ are the input and output grayscale value of the pixel at the location (i,j). a and b represent the minimum and maximum grayscale values before stretching, while c and d are the expected corresponding values after stretching. k is the adjustment coefficient, usually set to 1. According to the normal distribution, we take grayscale values in the range of 2σ for stretching, that is, keeping pixels with a distribution in the range of about 96%. As for pixels out the range of 2% - 98%, we set them the grayscale value of 2% or 98% pixels. To further increase the stability and reliability of features in low texture and contrast parts, a multi-feature strategy is adopted. Single feature may not output satisfactory results in remote sensing images. While different feature extraction algorithms can complement each other and help to solve this problem. We mainly use three mixed features of SIFT(Lowe, 1999), SURF(Speeded Up Robust Features)(Bay et al., 2006), and ORB(Oriented FAST and Rotated BRIEF)(Rublee et al., 2011) in this paper. Compared with SURF and ORB, SIFT is more robust, resulting in a better performance in complex scenarios. While SURF and ORB are more efficient based on simpler algorithms. These three features are extracted independently and combined to get final feature points. After obtaining feature points, FLANN(Fast Library for Approximate Nearest Neighbors)(Muja, Lowe, 2009) is used for fast matching, and RANSAC(RANdom SAmple Consensus)(Fischler, Bolles, 1981) algorithm is used to eliminate the error matches. Finally, the least-squares method is used to solve the optimal transformation model. We select 2D Affine as the band-wise transformation model in this paper. In order to reduce the influence of grayscale difference between different bands on matching, we adopt "iterative reference" strategy based on the idea of adjacent bands are more similar in spectrum signatures. It takes the resampled output band of this iteration as the input for the next iteration.Some results in coarse stage are shown in Fig.3.

3.2 Fine Band Registration

The coarse stage eliminates large-scale translation, rotation, and scaling between bands, but it is not accurate enough. Results from the coarse stage are used as initial values for further optimization. Inspired by (Peng et al., 2012), the registration of hyperspectral bands can be converted to the optimization of the rank of a low-rank matrix. Suppose $I_1, I_2, ..., I_n$ are images aligned with each other, w and h are width and height of the image. Images can be reduced to n one-dimensional vector $v_1, v_2, ..., v_n$ by stacking rows or columns of every image together, the length of this vector is $m = w \times h$. Since images are aligned and similar in content, there will be a linear relationship
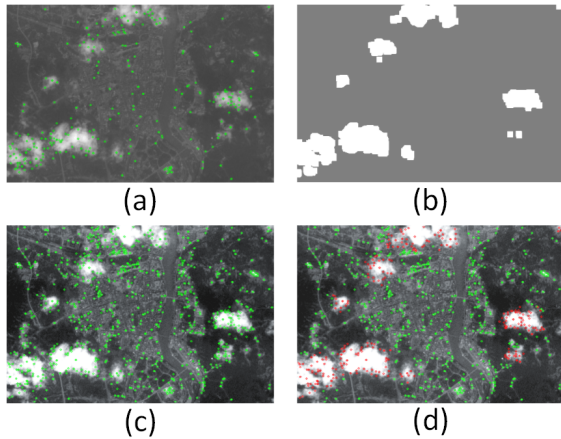
Figure 3. Results of the coarse stage. (a) Features extracted on the raw image. (b) Cloud mask of (a). (c) Features extracted on the stretched image using Eq.2. (d) Filtered result of (c) based on cloud mask (b) and RANSAC. Red and green points indicate cloud contour and ground points respectively.

among $v_1, v_2, ..., v_n$. Stack these vectors by columns to form a new low-rank matrix $A = [v_1, v_2, ..., v_n]$. The rank of A will be minimum when this linear relationship is strongest, which means that images are aligned.

Considering practical multi/hyperspectral bands of remote sensing images, this linear relationship is destroyed to a certain extent due to cloud movement, occlusion, and spectral reflection characteristics of different ground objects between bands. The aforementioned method cannot be used directly and it is necessary to eliminate the non-linear change of grayscale values between bands as much as possible. Inspired by RPCA(Robust Principal Component Analysis)(Candès et al., 2011), it can be resolved by decomposing the band into a low-rank part and a sparse part. Suppose $B_1, B_2, ..., B_n$ are bands in multi/hyperspectral image, B_i indicates the i-th band$(i \leq n)$. These multi/hyperspectral bands can be expressed as: $B_1 = B_1^l + B_1^s$, $B_2 = B_2^l + B_2^s$,..., $B_n = B_n^l + B_n^s$, while B_i^l indicates the low-rank part of i-th band, and B_i^s indicates the sparse part. Generally, the value of B_i^s is large but only affects a part of the band, and is used to correct the non-linear variation of grayscale values between bands. Let $vec()$ be the operation that converting a two-dimensional band into a one-dimensional vector by rows or columns and D be the matrix after each band is stacked, so we can write Eq.3.

$$D = [vec(B_1)|vec(B_2)|\cdots|vec(B_n)] = L + S \quad (3)$$

Where $L = [vec(B_1^l)|vec(B_2^l)|\cdots|vec(B_n^l)]$ is a low-rank matrix used to describe the overall linear relationship between bands. $S = [vec(B_1^s)|vec(B_2^s)|\cdots|vec(B_n^s)]$ is a sparse matrix used to express the non-linear relationship components between bands, such as the non-linear grayscale change caused by spectral characteristics of the ground objects. In addition to the non-linear changes in grayscale between bands, there is some misalignment which breaks the linear relationship between bands. For $B_1, B_2, ..., B_n$ bands, we can introduce n transformations $\tau = \{\tau_1, \tau_2, \cdots, \tau_n\}$ (such as affine, homography, etc.) to make $B_1 \circ \tau_1, B_2 \circ \tau_2, \cdots, B_n \circ \tau_n$ aligned as written Eq.4 and adjust transformations τ simultaneously during the optimization for finding the minimum rank of D.

$$D \circ \tau = [vec(B_1 \circ \tau_1)|vec(B_2 \circ \tau_1)|\cdots|vec(B_n \circ \tau_1)] \quad (4)$$

Both the non-linear changes of grayscale and misalignment between bands can be expressed by Eq.4, these two variables can be optimized at the same time to achieve accurate registration between bands. Our goal is to find the optimal low-rank matrix L and transformation τ. If bands are strictly aligned, the rank of L should be optimized as much as possible, and the objective function is as follows:

$$\min_{L,S,\tau} rank(L) \quad s.t. \quad D \circ \tau = L + S, \|S\|_0 < k \quad (5)$$

The L_0 norm indicates the number of non-zero elements in the sparse matrix S. The above equation can be more conveniently solved in the Lagrangian form:

$$\min_{L,S,\tau} rank(L) + \gamma\|S\|_0 \quad s.t. \quad D \circ \tau = L + S \quad (6)$$

Where γ is a parameter used to adjust the weight. While in Eq.6 both the objective and constraints are nonconvex and non-linear, it is hard to solve this optimization problem directly. To solve this problem, we adopt the methods in (Chandrasekaran et al., 2011),(Lin et al., 2009),(Peng et al., 2012). We convert the nonconvex objective function in Eq.6 to its convex surrogate constraint, replace the $rank()$ with the nuclear norm or sum of the singular values: $\|L\|_* = \sum_{i=1}^m \sigma_i(L)$, where m is the row number of L and then replace the L_0 norm $\|S\|_0$ with L_1 norm:$\sum_{ij} |S_{ij}|$. After applying this operation to Eq. 6, we can get a new objective function to optimize in Eq.7.

$$\min_{L,S,\tau} \|L\|_* + \lambda\|S\|_1 \quad s.t. \quad D \circ \tau = L + S \quad (7)$$

Here λ is a coefficient. Suggested by (Candès et al., 2011), it should be $\sqrt{C}/m$, where C is typically a constant and usually set to $C \approx 1$. m is the row number of L and equals to the product of w and h of a certain band. Now the objective function is continuous and convex. The only problem is the non-linearity of constraint $D \circ \tau = L + S$. If τ changes very small, we can linearize the current estimate of τ to approximate this constraint. We assume that transformation $\tau = [\tau_1|\tau_2|\cdots|\tau_n] \in R^{p \times n}$, p suggests the number of parameters for the transformation model. For example, $p = 6$ in 2D Affine model. n indicates the number of bands needs to be aligned. For $\Delta\tau \in R^{p \times n}$, we have $D \circ (\tau + \Delta\tau) \approx D \circ \tau + \sum_{i=1}^n J_i \Delta\tau_i \varepsilon_i^T$, where $J_i \approx \frac{\partial}{\partial\zeta} vec(B_i \circ \zeta)|_{\zeta=\tau_i} \in R^{p \times n}$ is the Jacobian of the i-th band corresponding to the transformation parameters τ_i and $\{\varepsilon_i\}$ is the standard basis for R^n. Eq.7 can be written as follows:

$$\min_{L,S,\Delta\tau} \|L\|_* + \lambda\|S\|_1 \quad s.t. \quad D \circ \tau + \sum_{i=1}^n J_i \Delta\tau_i \varepsilon_i^T = L + S \quad (8)$$

Here Eq.8 is a convex optimization problem, where $L, S, \Delta\tau$ are unknowns. By continuously iterating to solve Eq.8, we can get a series of convex problems until it converges. Finally, we can get the minimum estimate result of Eq.5.

4. EXTEXPERIMENTS

4.1 Data Introduction

The experiment data comes from Zhuhai-1 hyperspectral satellite constellation(OHS-A,B,C,D), which was launched on April 26, 2018. It is the first commercial hyperspectral satellite in China and opens a new era of commercialization(Orbita, 2020). A push-broom mode is adopted for hyperspectral imaging in Zhuhai-1 satellite. The imaging system is composed of three separated CMOS(Complementary Metal Oxide Semiconductors) sensors with an overlap of fifty pixels among them(Jiang et al., 2019). Every CMOS sensor contains 5056×2968 pixels with a GSD(Ground Sample Distance) of 10m, and is divided into 32 stripe regions along the flight direction with a filter on its surface, covering a range of 400 - 1000nm in spectrum. In practice, 8 lines are used for integral imaging, as shown in Fig.4.
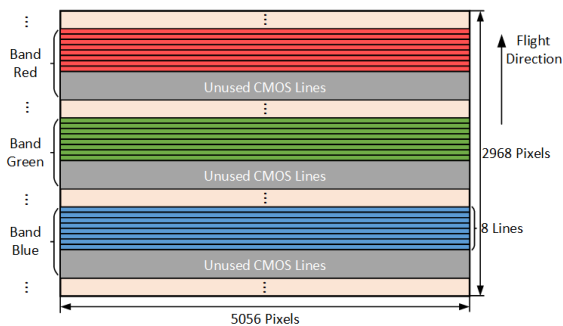

Figure 4. One CMOS sensor in Zhuhai-1 imaging system. Each band is captured by eight lines with a different location.

The different imaging position on CMOS leads to the misalignment between bands. However, we cannot directly eliminate it by history data due to its variation among images with different time and sensors, which is caused by random factors during the imaging. For Zhuhai-1 Satellite, the difference between adjacent bands along the flight direction in the raw image is 64.6 pixels in average, and the cumulative offset of all bands reaches about 1950 pixels in general. Experiments were performed on a total of 102 hyperspectral images, of which 83 cloudy images and 19 cloud-free images using all 32 bands with a size of 600×400 pixels, as listed in Tab.1.

Data	Satellite	Number	Size(pixels)
Cloudy Images	OHS-A	16	
	OHS-C	57	
	OHS-D	10	600×400
Cloud-free Images	OHS-A	7	
	OHS-C	12	

Table 1. Information about experiment data. We did not use data of OHS-B because we did not get data from it.

4.2 Evaluation Method

Registration accuracy can be evaluated by manually selecting tie points on different bands. However, it is not realistic to use this method with a large amount of data in experiments. (Zhu et al., 2013) uses automatic feature extraction and matching to find tie points for evaluation in multispectral images. However, almost no ground object appears in all hyperspectral bands due to diverse reflection characteristics, which brings great challenges for tie point extraction. Although the aforementioned "iterative reference" strategy can handle some spectral variation between bands, it is not feasible to use it to obtain the same tie point on

all bands directly. But we can get the band-wise transformation by feature extraction and matching first, and then multiply all band-wise transformations together to get an overall transformation for evaluation. These band-wise transformations have an approximate direction due to the unique push-broom imaging, which means the overall transformation reflects the maximum registration accuracy. In other words, the accuracy for any two bands between the first and the last band in the aligned image will be smaller than it. For one checkpoint, we calculate Euclidean distance between its two corresponding coordinates P in the first band and P' in the last band as the registration error. For one image, we select some evenly distributed points and then transform them to the last band based on the overall transformation. The final registration accuracy of this image is calculated by RMSE(Root Mean Squared Error) of all selected checkpoints, as written in Eq.9.

$$Acc_{reg} = \sqrt{\frac{\sum_{i=1}^{n} err_i}{n}} \qquad (9)$$

Where err_i is the registration error for i-th checkpoint, n is the number of checkpoints. We extract 4000 SIFT feature points for band-wise 2D Affine transformations and select 17 checkpoints for evaluation, as shown in Fig.5.
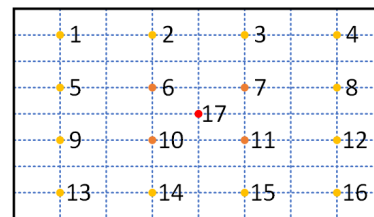

Figure 5. The map of checkpoint distributions. 16 checkpoints are selected from every other grid in 8×8 grids. Finally including the center point, a total of 17 checkpoints.

4.3 Results and Comparisons

Because we do not get operational parameters for building imaging model, only vision-based methods are compared. To verify our coarse-to-fine registration framework, "direct method", "cloud mask method" and our method were performed on 83 cloudy images respectively. "direct method" means we extract and match features directly and use it for building transformations without filtering any cloud-contour points. "cloud mask method" refers that we take into account the movement of clouds and use methods proposed in this paper to eliminate the movement and filter tie points. As aforementioned, a mix-feature strategy(SIFT, SURF, and ORB) is used for feature extraction. We extract a maximum number of 20,000 feature points per band. And the morphological operation kernel size for generating the cloud mask is 7, the number of iterations is 8. Finally, we calculate the RMSE of all 17 checkpoints in every image as the registration error using Eq.9. The registration accuracy of all 83 cloudy images is shown in Fig.6. The comparison of some registration results in cloudy images is shown in Fig.7. It can be seen that the "direct method" has a better registration result on cloud contours, while ghosting appears on ground objects. "cloud mask method" has partially eliminated the influence of cloud movement, making the registration of the ground objects slightly better. Compared with the first

two methods, our method can best eliminate the influence of cloud movement and achieve the highest registration accuracy for ground objects. Comprehensive statistics are performed on all 83 cloudy images, results are shown in Tab.2. In cloudy images, the mean RMSE of our method on all test data is 0.6 pixels. Compared with the 3.44 pixels accuracy of the "direct method", it is significantly improved by about 80%. To evaluate the variation of accuracy, we compare the biggest and smallest RMSE among all separated bands. Our method achieves an RMSE within a range of 0.28 - 0.9 in pixel, which means all bands achieved an accuracy better than 1 pixel. The comparisons of the first and the last band of the same ground features are shown in Fig.8. And the composite images after band registration are shown in Fig.9. In order to test the generality of our method, experiments were also performed on 19 cloud-free images. Since there are no moving clouds in images, the "direct method" is compared with our method. The results are shown in Tab.3. For images without clouds, the mean RMSE registration accuracy of the "direct method" is 0.81 pixels while we can continue improving about 50% to a mean RMSE of 0.41 pixels on test data. Some registration results are shown in Fig.10.
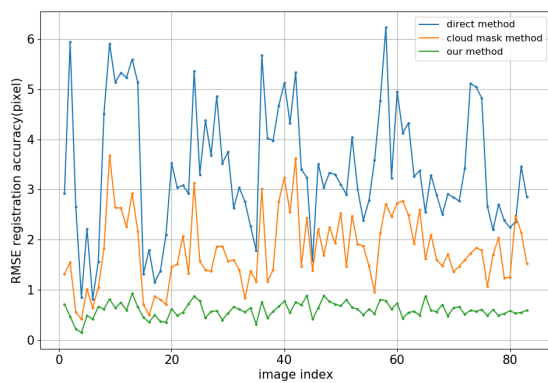

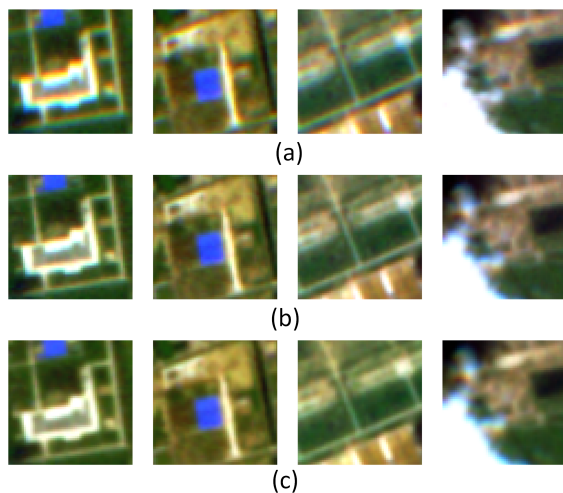Figure 6. RMSE registration accuracy of "direct method", "cloud mask method" and our method in 83 cloudy images.

Figure 7. Registration result (band 2(blue),7(green),14(red)) comparison of different methods in cloudy images. (a) are results of the "direct method". (b) are results of the "cloud mask method". (c) are results of our method.

After a large number of experiments, our method can achieve an accuracy of mean RMSE of 0.6 pixels in cloudy images and 0.41 pixels in cloud-free images. It is slightly better than res-
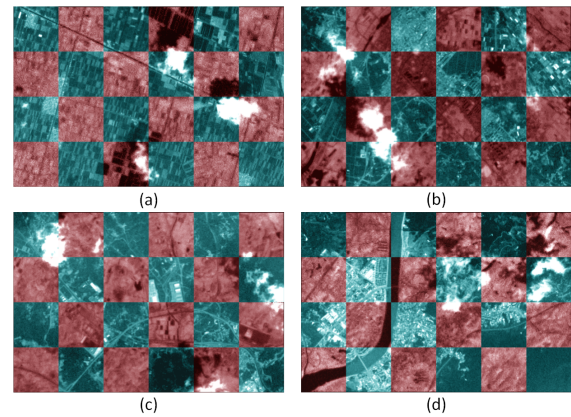

Figure 8. Comparison of the first band(marked blue) and the last band(marked red) in cloudy images using our framework.(a) is the farmland.(b) is the city.(c) is the road.(d) is the river.
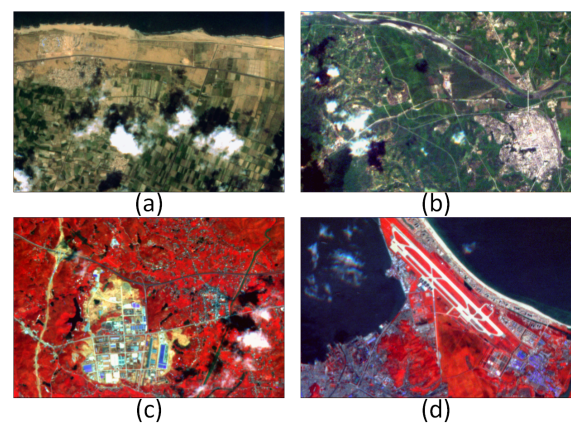

Figure 9. Color composite images of our methods in cloudy images. (a), (b) are true-color images with band 2,7,14. (c), (d) are pseudocolor images with band 2,7,21.
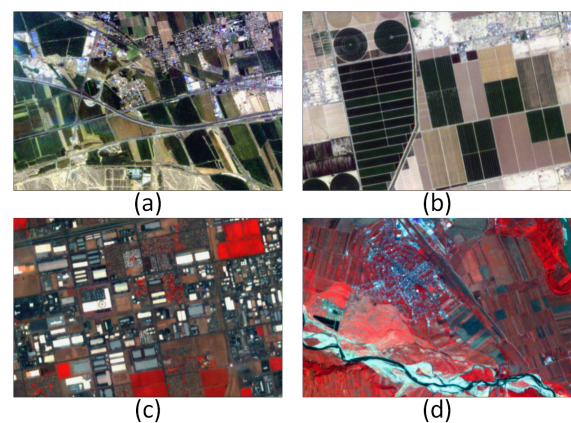

Figure 10. Color composite results of our methods in cloud-free images. (a), (b) are true-color images with band 2,7,14. (c), (d) are pseudocolor images with band 2,7,21.

ults in (Jiang et al., 2019), which uses the geometric positioning model for Zhuhai-1 hyperspectral satellite band registration. And we also simplifies the production of data to some extent. In addition, based on our framework, every band is decomposed into a sparse and a low-rank part.We can effectively identity and remove speckle noises and unevenly stripes using sparse mat-

Criterion	Max dx	Min dx	RMSE dx	Var. dx	Max dy	Min dy	RMSE dy	Var. dy	Max ds	Min ds	RMSE ds	Var. ds
Direct Method	1.527	0.420	1.013	0.145	4.148	2.168	3.223	0.474	4.414	2.377	3.435	0.499
Cloud Mask Method	1.321	0.313	0.850	0.120	2.229	0.661	1.490	0.301	2.602	0.971	1.796	0.304
Our Framework	**0.601**	**0.097**	**0.364**	**0.031**	**0.736**	**0.125**	**0.439**	**0.043**	**0.901**	**0.284**	**0.596**	**0.041**
Increase PCT1	13.48	25.40	16.06	17.60	46.25	69.52	53.79	36.49	41.05	59.13	47.72	39.10
Increase PCT2	54.53	69.15	57.18	74.16	66.96	81.17	70.56	85.72	65.39	70.76	66.80	86.51
Increase PCT3	**60.66**	**76.99**	**64.06**	**78.71**	**82.24**	**94.26**	**86.40**	**90.93**	**79.60**	**88.05**	**82.64**	**91.78**

Table 2. Comparison of "direct method", "cloud mask method" and our framework in 83 cloudy images. (Accuracy unit: pixel). Max dx, dy and ds are the max errors among all images in x-direction, y-direction and plane respectively, while Min dx, dy and ds are the opposite. Var. means the variance. Increase PCT1,PCT2,PCT3 are the percentage increase in "cloud mask method" relative to the "direct method",our method relative to the "cloud mask method", and our method relative to the "direct method" respectively.

Criterion	Max dx	Min dx	RMSE dx	Var. dx	Max dy	Min dy	RMSE dy	Var. dy	Max ds	Min ds	RMSE ds	Var. ds
Direct Method	0.671	0.078	0.390	0.043	1.060	0.252	0.663	0.084	1.263	0.366	0.806	0.087
Our Framework	**0.467**	**0.003**	**0.242**	**0.021**	**0.547**	**0.072**	**0.315**	**0.026**	**0.689**	**0.150**	**0.412**	**0.031**
Increase PCT	**30.46**	**61.84**	**37.89**	**51.03**	**48.39**	**71.32**	**52.50**	**68.81**	**45.46**	**59.03**	**48.92**	**65.04**

Table 3. Comparison of "direct method" and our framework in 19 cloud-free images. (Accuracy unit: pixel). Max dx, dy and ds are the max errors among all images in x-direction, y-direction and plane respectively, while Min dx, dy and ds are the opposite. Var. means the variance. Increase PCT is the percentage increase in our framework relative to the "direct method".

rix, while keeping spectral signatures of ground objects. And inpainting corrupted or damaged parts by low-rank component of that band.Some results are shown in Fig.11.
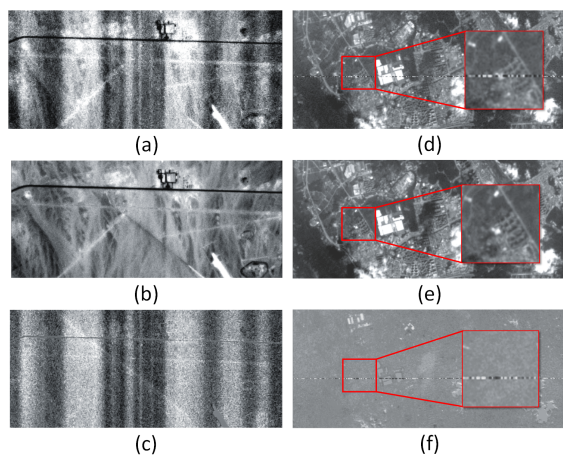

Figure 11. Denoising results using our method. (a) Original band with very large uneven stripe noise. (b) Aligned and noise-free band. (c) Separated stripe noise. (d) Original band with the damaged line on the lower middle. (e) Aligned and inpainted band. (f) Separated noise and damaged line.

5. CONCLUSION

In this paper, we analyzed the influence of moving clouds on multi/hyperspectral band registration and proposed a novel coarse-to-fine framework, which achieves high-accuracy in cloudy images. In the coarse stage, a registration is perfumed based on generated cloud mask and feature matching. In the refinement stage, the low-rank analysis and RASL method are used to optimize the rank of coarse results for fine registration results. After a large number of experiments, our method can achieve a good accuracy in both cloudy and cloud-free images, which can meet the needs of subsequent applications. In the future, data from other hyperspectral satellites can also be used for testing. In addition, how to give a more accurate initial value for optimization is also an open problem needs to be studied.

REFERENCES

Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., Gumley, L. E., 1998. Discriminating clear sky from clouds with MODIS. *Journal of Geophysical Research: Atmospheres*, 103(D24), 32141–32157.

Bay, H., Tuytelaars, T., Van Gool, L., 2006. Surf: Speeded up robust features. *2006 European conference on computer vision(ECCV)*, Springer, 404–417.

Candès, E. J., Li, X., Ma, Y., Wright, J., 2011. Robust principal component analysis? *Journal of the ACM (JACM)*, 58(3), 1–37.

Chandrasekaran, V., Sanghavi, S., Parrilo, P. A., Willsky, A. S., 2011. Rank-sparsity incoherence for matrix decomposition. *SIAM Journal on Optimization*, 21(2), 572–596.

Deng, J., Wang, H., Ma, J., 2016. An automatic cloud detection algorithm for landsat remote sensing image. *2016 4th International Workshop on Earth Observation and Remote Sensing Applications (EORSA)*, IEEE, 395–399.

Dong, G., Khim, N. T., Kim, P. S., Adrian, Y. C. L., 2011. Dense registration of push broom hyperspectral imageries. *2011 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, IEEE, 1–4.

Fischler, M. A., Bolles, R. C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381–395.

Hagolle, O., Huc, M., Pascual, D. V., Dedieu, G., 2010. A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENμS, LANDSAT and SENTINEL-2 images. *Remote Sensing of Environment*, 114(8), 1747–1755.

Harris, C. G., Stephens, M. et al., 1988. A combined corner and edge detector. *Alvey vision conference*, 15, Citeseer, 10–5244.

Hu, T., Zhang, H., Shen, H., Zhang, L., 2014. Robust registration by rank minimization for multiangle hyper/multispectral remotely sensed imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6), 2443–2457.

Jedlovec, G. J. et al., 2008. Spatial and temporal varying thresholds for cloud detection in GOES imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 46(6), 1705–1717.

Jiang, Y., Wang, J., Zhang, L., Zhang, G., Li, X., Wu, J., 2019. Geometric Processing and Accuracy Verification of Zhuhai-1 Hyperspectral Satellites. *Remote Sensing*, 11(9), 996.

Jiang, Y., Zhang, G., TANG, X., Zhu, X., Huang, W., Pan, H., Qin, Q., 2013. Research on the high-accuracy band-to-band registration method of ZY-3 multispectral image. *Acta Geodaetica et Cartographica Sinica*, 42(6), 884–890.

King, M. D., Platnick, S., Menzel, W. P., Ackerman et al., 2013. Spatial and temporal distribution of clouds observed by MODIS onboard the Terra and Aqua satellites. *IEEE Transactions on Geoscience and Remote Sensing*, 51(7), 3826–3852.

Li, X., Wang, L., Cheng, Q., Wu, P., Gan, W., Fang, L., 2019a. Cloud removal in remote sensing images using nonnegative matrix factorization and error correction. *ISPRS journal of photogrammetry and remote sensing*, 148, 103–113.

Li, Z., Shen, H., Cheng, Q., Liu, Y. et al., 2019b. Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors. *ISPRS Journal of Photogrammetry and Remote Sensing*, 150, 197–212.

Lin, Z., Ganesh, A., Wright, J., Wu, L., Chen, M., Ma, Y., 2009. Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix. *Coordinated Science Laboratory Report no. UILU-ENG-09-2214, DC-246*.

Lowe, D. G., 1999. Object recognition from local scale-invariant features. *Proceedings of the seventh IEEE international conference on computer vision*, 2, IEEE, 1150–1157.

Ma, W., Wen, Z., Wu, Y., Jiao, L., Gong, M., Zheng, Y., Liu, L., 2016. Remote sensing image registration with modified SIFT and enhanced feature matching. *IEEE Geoscience and Remote Sensing Letters*, 14(1), 3–7.

Muja, M., Lowe, D. G., 2009. Fast approximate nearest neighbors with automatic algorithm configuration. *VISAPP (1)*, 2(331-340), 2.

Orbita, 2020. Orbita. *https://www.myorbita.net/fzlc/list$_6$.aspx*.

Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1), 62–66.

Ozkan, S., Efendioglu, M., Demirpolat, C., 2018. Cloud detection from rgb color remote sensing images with deep pyramid networks. *2018 IEEE International Geoscience and Remote Sensing Symposium(IGARSS)*, IEEE, 6939–6942.

Pan, J., Zhu, Y., Wang, M. et al., 2011. Parallel band-to-band registration for hj-1a1b ccd images using openmp. *2011 International Symposium on Image and Data Fusion*, IEEE, 1–4.

Peng, Y., Ganesh, A., Wright, J., Xu, W., Ma, Y., 2012. RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images. *IEEE transactions on pattern analysis and machine intelligence*, 34(11), 2233–2246.

Randall, D. A., Corsetti, T. G., 1989. Earth radiation budget and cloudiness simulations with a general circulation model. *Journal of the atmospheric sciences*, 46(13), 1922–1942.

Rublee, E., Rabaud, V., Konolige, K., Bradski, G., 2011. Orb: An efficient alternative to sift or surf. *2011 International conference on computer vision(ICCV)*, IEEE, 2564–2571.

Shao, Z., Pan, Y., Diao, C., Cai, J., 2019. Cloud detection in remote sensing images based on multiscale features-convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 57(6), 4062–4076.

Wang, S., Quan, D., Liang, X., Ning, M., Guo, Y., Jiao, L., 2018. A deep learning framework for remote sensing image registration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145, 148–164.

Wang, T., Shi, J., Yan, G., Gao, B., Chen, L. et al., 2016. Toward a general method for detecting clouds and shadows in optical remote sensing imagery. *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, 585–588.

Wen, F., Zhang, Y. et al., 2018. Two-pass robust component analysis for cloud removal in satellite image sequence. *IEEE Geoscience and Remote Sensing Letters*, 15(7), 1090–1094.

Xu, M., Jia, X., Pickering, M., Jia, S., 2019. Thin cloud removal from optical remote sensing images using the noise-adjusted principal components transform. *ISPRS journal of photogrammetry and remote sensing*, 149, 215–225.

Ye, Y., Shan, J., 2014. A local descriptor based registration method for multispectral remote sensing images with non-linear intensity differences. *ISPRS Journal of Photogrammetry and Remote Sensing*, 90, 83–95.

Yi, Z., Zhiguo, C., Yang, X., 2008. Multi-spectral remote image registration based on SIFT. *Electronics Letters*, 44(2), 107–108.

Zhang, Y., Wen, F., Gao, Z., Ling, X., 2019. A Coarse-to-Fine Framework for Cloud Removal in Remote Sensing Image Sequence. *IEEE Transactions on Geoscience and Remote Sensing*, 57(8), 5963–5974.

Zhu, Y., Wang, M., Pan, J., 2013. An automatic accuracy evaluation approach of band registration for multi-spectral imagery. *2013 IEEE International Geoscience and Remote Sensing Symposium(IGARSS)*, IEEE, 1481–1484.

Zhu, Z., Wang, S., Woodcock, C. E., 2015. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, 159, 269–277.