

Automatic Stitching for Hyperspectral Images Using Robust Feature Matching and Elastic Warp

Yujie Zhang, Zhiying Wan, Xingyu Jiang^{ID}, and Xiaoguang Mei^{ID}

Abstract—Hyperspectral images, which contain not only spatial information but also rich spectral information, have been extensively applied to the fields of agriculture, urban planning, etc. However, it is difficult for a single image to cover a large area. Therefore, it requires to take photos of various parts and apply image stitching technology to obtain a panoramic hyperspectral image. When the viewpoint of the scene changes a lot, the ghost issue will occur with traditional methods. In order to get the high-precision resultant panoramas, this article proposes an automatic image stitching algorithm for hyperspectral images using robust feature matching and elastic warp. Our method contains two stages. The first stage is to choose one band as reference band and obtain the panorama in a single band. In particular, we extract feature points by scale-invariant feature transform. Then, we propose an efficient algorithm called multiscale top K rank preservation algorithm, for establishing robust point correspondences between two sets of points. Next, we adopt robust elastic warp to obtain the panorama of each band. The second stage is to stitch all remaining bands based on the transformation obtained in the first stage and fuse the information of all bands together to get the final panoramic hyperspectral image. Extensive experiments have demonstrated the effectiveness of our proposed method.

Index Terms—Elastic warp, feature matching, hyperspectral images, image stitching.

I. INTRODUCTION

RECENTLY, hyperspectral imaging technique has been extensively applied to a wide range of fields from remote sensing to biomedicine [1]. The spectral information provided by hyperspectral images is more abundant than RGB images thanks to the high spectral resolution and wider sensing spectrum [2]. The spectrum of hyperspectral images provides us a new way to achieve higher accuracy in the identification of land covers and the diagnosis of some diseases. With the development of unmanned aerial vehicles (UAVs), the application of hyperspectral images in remote sensing has been improved. However, to achieve high spatial resolution, the instantaneous field of view will be small. It is difficult to obtain information of a large area through a single image. Therefore, the study of image stitching for hyperspectral images is meaningful.

Generally, image stitching includes two steps, image matching and radiometric normalization. Image matching is typically

Manuscript received April 11, 2020; revised May 26, 2020; accepted June 4, 2020. Date of publication June 9, 2020; date of current version June 18, 2020. This work was supported by the National Natural Science Foundation of China under Grant 61903279. (*Corresponding author: Xiaoguang Mei.*)

The authors are with the Electronic Information School, Wuhan University, Wuhan 430072, China (e-mail: zhangyujie14159@outlook.com; WanZhiying@hotmail.com; jiangx.y@whu.edu.cn; meixiaoguang@gmail.com).

Digital Object Identifier 10.1109/JSTARS.2020.3001022

solved in a two-step manner [3], [4]. The common idea is to extracting a sparse set of features first, based on establishing correspondences between points, edges, corners, or other geometric entities [5]–[9]. The well-known scale-invariant feature transform (SIFT) [5] is very robust, and speeded up robust features (SURF) [6] improve the computation time of SIFT by using a fast local gradient computation. Spectral-spatial SIFT (SS-SIFT) extracts 3-D features from hyperspectral images, but the computation time is not feasible for real-time applications. To improve the registration accuracy between two images, it is required to remove the false matches after feature point extraction and matching. Random sampling consensus (RANSAC) [10], [11] is a representative method for eliminating false matches. But RANSAC can be sensitive to selecting the correct noise threshold that defines which data points fit a model instantiated with a certain set of parameters. On the basis of RANSAC, maximum likelihood estimation sample consensus [12] and progressive sample consensus [13] were proposed as effective variants. Although these methods achieve good results, they cannot be applied to all situations. When the motion of image scene is nonrigid, which cannot be characterized by a parametric model, these methods will be invalid. To solve this problem, the nonparametric fitting methods were introduced, such as vector field consensus (VFC) [14], [15] and locality preserving matching (LPM) [16], [17]. However, the results obtained by these methods are not perfect if there is a large proportion of outliers.

For image alignment, global deformation method AutoStitch was first introduced by Brown and Lowe [18]. It estimates an optimal global transformation matrix for each input image and uses the matrix to align all images, which is not parallax-tolerant. To achieve better stitching quality, methods based on space-varying transformation were proposed, such as as-projective-as-possible (APAP) [19] and shape-preserving-half-projective (SPHP) [20]. In these methods, the input images are warped according to the computed deformations over the meshed plane. APAP adopts local adaptive projective transformation and obtains good results in the overlapped region. But APAP is similar to global projection and still remains the problem of distortion. SPHP solves this problem by using projective transformations, which is gradually changed to a global similarity transformation across the image. Combining APAP with SPHP, better results can be obtained. Based on the aforementioned works, adaptive as-natural-as-possible (ANAP) [21] was proposed to reduce the negative impact brought by unnatural rotation. However, unnatural rotation and scaling still exist while stitching multiple

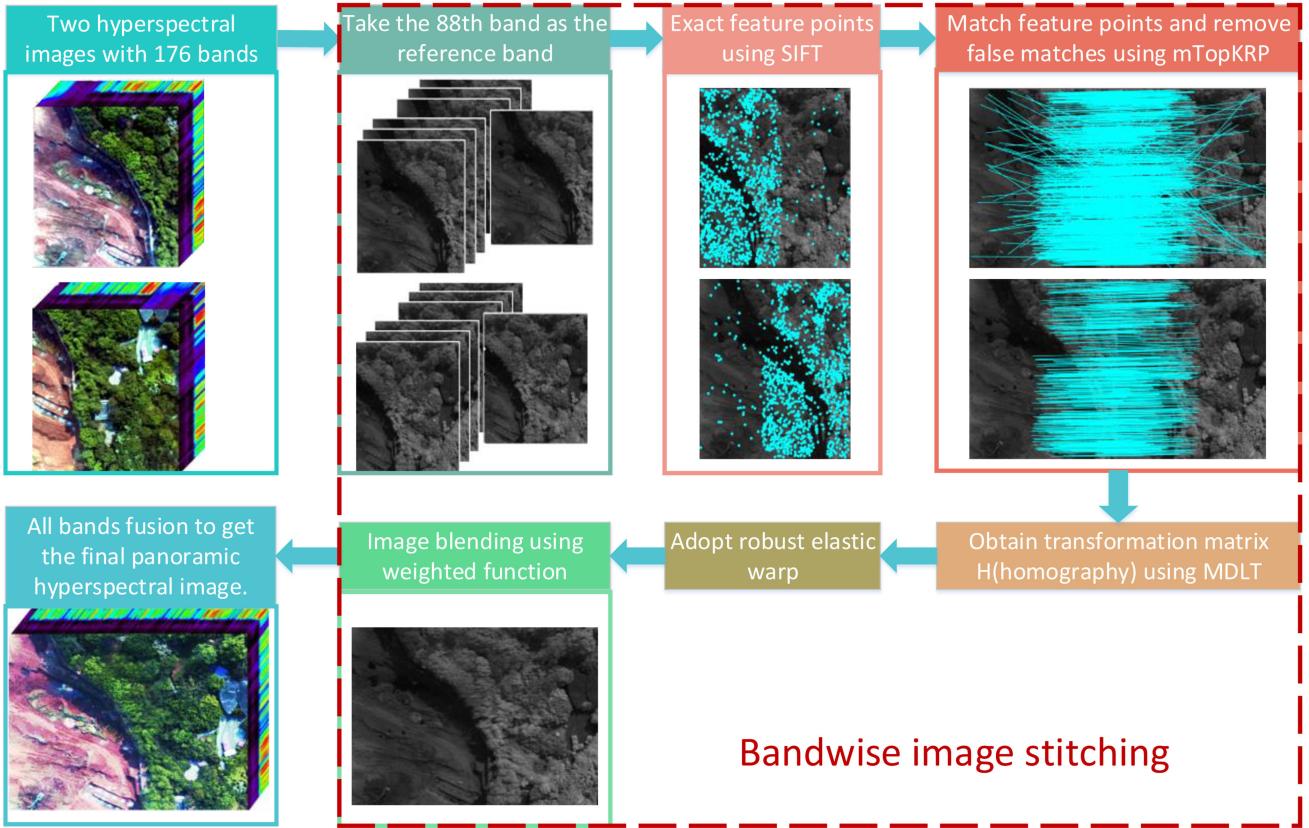


Fig. 1. Proposed flow of the hyperspectral image stitching.

images. A more robust stitching method natural image stitching with the global similarity prior (NISwGSP) was proposed by Chen and Chuang [22].

Although the stitching methods for RGB images have been greatly improved these years, they cannot be applied to hyperspectral images directly. Different from RGB images, each pixel of hyperspectral images has tens or hundreds of magnitude values corresponding to its wave bands. Therefore, the stitching algorithm should be efficient to avoid a large amount of calculation. For UAV-based hyperspectral images, the images may undergo scale change, parallax, unnatural rotation, or nonrigid transformation. The existing algorithm for hyperspectral images stitching proposed by Yuan *et al.* [23] is only applicable to medical hyperspectral images without parallax. Fang *et al.* [24] proposed their method by using the geometric data of images, which is limited and cannot be applied for complex scenes.

To solve the aforementioned challenge, we propose a robust method for hyperspectral images stitching according to the characteristics of hyperspectral remote sensing images, which is stable even when the images undergo scale change, rotation, or nonrigid transformation. We divide our work into two stages. The first one is to choose one band as reference band and obtain the panorama in a single band. We extract and match feature points by SIFT algorithm [5] and remove the false matches by our previously proposed multiscale Top K Rank Preservation (mTopKRP) [25] algorithm. Next, we construct robust elastic warp and reproject them to obtain seamless panorama in

greyscale. The method is parallax-tolerant and robust, which can deal with complex local distortion. The second stage is to stitch all the remaining bands using the matrix obtained from the first reference band and fuse all band information to obtain the final hyperspectral panorama.

This article is an extension of our previous work [25], and the main contributions contain two aspects. First, we introduce a new stitching strategy and apply it to hyperspectral images, which combines the mTopKRP algorithm and elastic warp to achieve high precision even when the scene suffers large view point changes. Second, we provide a more comprehensive evaluation of our method with more discussions, and the qualitative result analysis demonstrates the advantages of our method. The spectrum analysis of the stitched image shows the accurate results after the spectral matching.

II. METHODOLOGY

In this section, we introduce a new method for hyperspectral images stitching, including feature matching based on SIFT [5] and mTopKRP algorithm, as well as robust elastic warp and multibands fusion to obtain hyperspectral panorama. The flowchart of our method is shown in Fig. 1.

A. Image Stitching in a Single Band

The large number of bands in hyperspectral images leads to a significant increase in their data size. Therefore, directly

extracting feature points from hyperspectral data and stitching them will generate a large amount of calculation, which will take a lot of time and reduce the stitching efficiency. Due to that the image of each band contains the same spatial information, we choose one band as reference band and perform image stitching in a single band. We use the same transformation in the other bands to obtain panorama in all bands.

1) *Multiscale Top K Rank Preserving Matching*: In the gray image stitching, feature points extraction and matching are challenging due to local terrain changes, scale change, rotation, parallax, and nonrigid transformation existing in UAV-based hyperspectral images. To achieve scale, rotation, and illumination invariance, researchers usually use local descriptor to extract feature points, which leads to many false matches. Former algorithms for eliminating false matches typically use a parametric model and obtain unsatisfying results with large proportion of false matches. To solve the aforementioned challenge, the method based on mTopKRP [25] is introduced to eliminate the false matches, which can perform well in the aforementioned tough situations.

First, we use SIFT algorithm to extract feature points in images I_1 and I_2 . Then, we can obtain a set of N putative matches $S = \{(x_i, y_i)\}_{i=1}^N$ where x_i and y_i are the pixel coordinates. In order to get reliable correspondence, we need to remove the false matches precisely and acquire an optimal inlier set. We denote the optimal inlier set as R^* . The optimal solution is

$$R^* = \arg \min C(R; S, \lambda). \quad (1)$$

Through observation, the relationship among the neighboring feature points in a remote sensing image is usually stable and only changed slightly, even when the image undergoes scale change, rotation, or nonrigid transformation. According to mTopKRP, the similarity of the feature points' K-NN is used to measure whether the feature points are correctly matched and the cost function C is defined as

$$C(\mathbf{p}; S, \lambda) = \sum_{i=1}^N p_i (c_i - \lambda) + \lambda N \quad (2)$$

where $c_i = \frac{1}{M} \sum_{m=1}^M D_{km}(\sigma(x_i), \sigma(y_i))$, vector \mathbf{p} is associated with S , $p_i \in (0, 1)$. $p_i = 1$ denotes that the match is correct and this feature point is in inlier sets. On the contrary, $p_i = 0$ denotes that the match is wrong and this feature point is in outlier sets. $D_{km}(\sigma(x_i), \sigma(y_i))$ denotes the distance between the top K ranking lists of the two feature points' K-NN.

The cost function also applies a multiscale strategy, which measures the local K neighborhoods ranking similarity and calculates the cost function under different scales of K . We define a set of different K with $K = \{K_m\}_{m=1}^M$. The ideal situation is that c_i should be zero. Therefore, our cost function tries to obtain the maximum inlier number and keep the cost value to a minimum at the same time.

Given a putative match, the K-NN of all feature points can be calculated. Next, all the cost values of $\{c_i\}_{i=1}^N$ can be calculated. It is easy to observe that any putative match with a cost smaller than the parameter λ will lead to a negative term and decrease the objective cost function. Therefore, the optimal solution can

be defined by the following criterion:

$$p_i = \begin{cases} 1, & c_i \leq \lambda \\ 0, & c_i > \lambda \end{cases}. \quad (3)$$

In this case, the optimal inlier set R^* can be determined by

$$R^* = \{i | p_i = 1, i = 1, 2, \dots, N\}. \quad (4)$$

2) *Construct Robust Elastic Warps*: We have obtained the optimal inlier set of input images through mTopKRP, the next step is to calculate the homography matrices between two adjacent images corresponding to the input order. After local homography is obtained, the image I_1 can be projected onto the image I_2 to achieve alignment of the image. However, there are certain drawbacks associated with the use of projective transformation, such as shape distortion. Since the hyperspectral images to be stitched were taken by UAV, the shooting height is not high enough to avoid parallax. Some parts of the stitching images are severely stretched and magnified unevenly. Besides, the mosaic looks ghosted. Thus, we adopt the robust elastic warps to avoid projective distortion (stretched shapes and enlarged sizes). A homography \mathbf{H} is estimated, as shown in (5). Suppose the matched feature points: $\mathbf{p} = [x, y, 1]^T$ and $\mathbf{q} = [u, v, 1]^T$, then we have

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (5)$$

where \sim denotes equality up to a scale factor. The transformation of formula (5) can be expressed as

$$\mathbf{0}_{3 \times 1} = \begin{bmatrix} \mathbf{0}_{1 \times 3} & -\mathbf{p}^T & v\mathbf{p}^T \\ \mathbf{p}^T & \mathbf{0}_{1 \times 3} & -u\mathbf{p}^T \\ -v\mathbf{p}^T & u\mathbf{p}^T & \mathbf{0}_{1 \times 3} \end{bmatrix} \cdot \mathbf{h} = \mathbf{A}\mathbf{h}, \|\mathbf{h}\|^2 = 1 \quad (6)$$

where $\mathbf{h} = [h_{11}, h_{12}, \dots, h_{33}]^T$, $\mathbf{h} \in R^{9 \times 1}$ and $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N]^T$, $\mathbf{a}_i \in R^{2 \times 9}$, $i = 1, 2, 3, \dots, N$.

Then, we formulate (6) by DLT as

$$\mathbf{h} = \arg \min_{\mathbf{h}} \sum_{i=1}^N \|\mathbf{a}_i \mathbf{h}\|^2 = \arg \min_{\mathbf{h}} \|\mathbf{A}\mathbf{h}\|^2. \quad (7)$$

The image I_1 is divided into $C_1 \times C_2$ grids, and the coordinates of the grid vertices in the homogeneous coordinate system are \mathbf{p}_* . The local homography at arbitrary position \mathbf{p}_* in I_1 can be calculated by the MDLT [19] as

$$\mathbf{h}_* = \arg \min_{\mathbf{h}} \sum_{i=1}^N \|\omega_*^i a_i \mathbf{h}\|^2 = \arg \min_{\mathbf{h}} \|W_* \mathbf{A}\mathbf{h}\|^2 \quad (8)$$

where $\omega_*^i = \max(\exp(-\|\mathbf{p}_* - \mathbf{p}_i\|^2 / \sigma^2), \lambda)$, σ is a scale parameter and $\lambda \in [0, 1]$, $W_* = \text{diag}([\omega_*^1, \omega_*^1, \omega_*^2, \omega_*^2, \dots, \omega_*^N, \omega_*^N])$.

Equation (8) can be solved by the smallest right singular vector of $W_* \mathbf{A}$ [26]. However, the distortion problems can be caused by extrapolating the deformation inclination to the nonoverlapping region [27], [28]. Li *et al.* [29] linearly reduced the deformation function $g(x, y)$ to zero across the nonoverlapping region. This method reduces the distortion of the image outside the overlapping region to some extent, but the image



Fig. 2. Scenes for stitching were gathered by GaiaSky-mini sensor provided by SiChuan Dualix Spectral Imaging Technology Company, Ltd., which is a UAV-borne hyperspectral sensor whose wavelength ranges from 400 to 1000 nm. We synthesize the pseudocolor image by selecting the 92th (700.2 nm), 47th (547.6 nm), and 13th (436.5 nm) bands.

will still appear obvious distortion in our experiment. One way to overcome this problem is to use weighted optimization algorithm [30].

The image I_2 needs to be mapped to the image I_1 by the homography matrix. The weighted deformation function in the x -direction of the image I_2 coordinates can be expressed as

$$g_\mu(x, y) = \mu g(H(x, y)), (0 < \mu(x) < 1). \quad (9)$$

The value of μ linearly decrements from 1 to 0 from the left to the right of the image. The weighted deformation function in the y -direction changes in the same way, expressed as

$$h_\mu(x, y) = \mu h(H(x, y)), (0 < \mu(x) < 1). \quad (10)$$

According to the value of $\mu(x)$, the region with large shape variable of image I_2 has smaller deformation for image I_1 . Therefore, the distortion in overlapping boundary region is small. Next, we combine our warp with the global similarity transformation. After the refined point matches are received from the mTopKRP, we estimate the global similarity transformation \mathbf{S} by

$$J_A = \sum_{i=1}^n \|\mathbf{Sp}_i - \mathbf{q}_i\|^2 \quad (11)$$

where $\mathbf{S} = \begin{pmatrix} c & -s & t_x \\ s & c & t_y \end{pmatrix}$. $\{\mathbf{p}_i, \mathbf{q}_i\}, i = 1, \dots, n$ are the point matches.

The application of similar transformations has largely preserved the original perspective and mitigated distortion. We mix the homography and the similarity transformation using

the formula as

$$\mathbf{H}_q = \mu_h \mathbf{H} + \mu_s \mathbf{H}_s \quad (12)$$

with μ_h from 0 to 1 across the source image I_1 . The target image I_2 is transformed as

$$\mathbf{H}_p = \mathbf{H}_q \mathbf{H}^{-1}. \quad (13)$$

3) Linear Blending: We obtain the robust elastic warp to unify the image pixels to be stitched into one coordinating system. In the standard system, the transformed coordinates may not be integer values, so it is necessary to use the weighted method to perform gray-scale interpolation in the overlapped areas of the image.

Our method requires all the images to be input in order. Using the weighted function (linear) can merge many images into one image, and realize the linear blending of the all input images according to the proportion of each image.

The key point is to match the spectral values of the pixel points in the overlapped region. Due to that the spectral values of corresponding points of different bands in the overlapping area are different, we need to find out the relationship between the spectral value of the corresponding points on the same band. In this article, we composite images by linear blending [31]. The main idea is to generate weight map for each image. Typically, the weight at center is large, whereas the weight at edge is small. Each output pixel is a weighted average of inputs. This algorithm can effectively get seamless stitched images.

TABLE I
COMPARISON OF SIFT AND OTHER FEATURE POINTS DETECTION ALGORITHMS

Algorithm	Detected points	Matching points	Run time
SIFT [5]	6,688	1,176	10.12s
SURF [6]	467	80	0.849s
SS-SIFT [7]	8,304	82	950.311s

Algorithm 1: Automatic Stitching for Hyperspectral Images Using Robust Feature Matching and Elastic Warp.

Input: Hyperspectral images $I_i, i = 1, \dots, K$, each of which have M bands.

Output: Panoramic hyperspectral image without parallax

- 1: Choose the $\lfloor M/2 \rfloor$ th band as reference band
- 2: In the reference band, extract feature points by SIFT and obtain a putative matching set $S = \{(x_i, y_i)\}_{i=1}^N$
- 3: Construct ranking lists of each feature points based on multiscale K-NN
- 4: Calculate the cost value $\{c_i\}_{i=1}^N$ using (6) and obtain the inlier sets R^*
- 5: **for** $i = 1$ to $K - 1$ **do**
- 6: $j = i + 1$
- 7: Calculate the homography \mathbf{H} and similarity \mathbf{H}_s using (7)
- 8: Define nonoverlapping region and calculate the global projectivity by (9) and (10)
- 9: Obtain the global warp by mixing the \mathbf{H} and the \mathbf{H}_s using (12)
- 10: **end for**
- 11: Map the warped images $I_i, i = 1, \dots, K$ by (14) and (15)
- 12: Stitch the panorama using the same global warp for the remaining $(M - 1)$ bands
- 13: Linear fusion of all band spectral information

For the input warped images $I_i, i = 1, 2, \dots, K$, the stitched image can be expressed as

$$I = \frac{\sum_{i=1}^K w_i I_i}{\sum_{i=1}^K w_i} \quad (14)$$

where w_i is a weighted function of I_i , and it can be calculated as

$$w_i(x, y) = 1 - \frac{d_i(x, y)}{d_{i,\max}} \quad (15)$$

where $d_i(x, y)$ is the distance from point (x, y) to the center of I_i , and $d_{i,\max}$ is the maximum distance from the point of I_i where the pixel value is not 0 to the center point.

B. All Bands Fusion

In this section, we use the robust elastic warp of one band as the global warp to stitch the panorama of other remaining bands, and then we can get all band spatial information. We choose two known pairs of corresponding points. For the single-band image stitching, the two images have the same wavelength, but the spectral values are different. Then, the spectral values of all

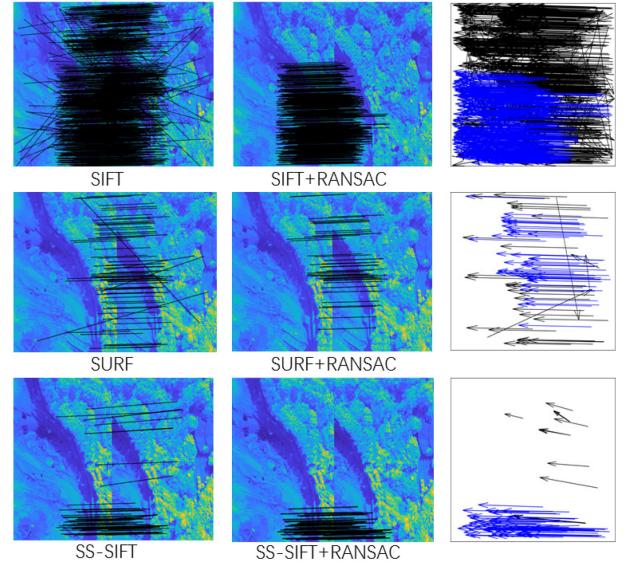


Fig. 3. Feature point matching of SIFT [5], SURF [6], and SS-SIFT [7], and false matches removing by RANSAC: 1254 tentative matches using SIFT and 557(44.42%) inlier matches using RANSAC; 80 tentative matches using SURF and 49(61.25%) inlier matches using RANSAC; and 82 tentative matches using SS-SIFT and 64(78.04%) inlier matches using RANSAC (top to bottom and left to right). Black lines connect putative matching points. Blue ones connect correct matching points.

TABLE II
COMPARISON OF THE EVALUATION INDICATORS AMONG SIX ALGORITHMS

algorithm	EOG	DFT	Variance
ANAP [21]	2.577×10^8	3.194×10^9	3.172×10^6
APAP [19]	2.646×10^8	3.177×10^9	3.230×10^6
DHW [32]	2.606×10^8	3.192×10^9	3.209×10^6
NISwGSP [22]	2.579×10^8	3.193×10^9	3.204×10^6
ELA [29]	2.573×10^8	2.877×10^9	3.152×10^6
Our Warp	2.776×10^8	4.750×10^9	3.742×10^6

points in the overlapping region of this band are calculated using formula (14). After the single band spectral stitching, both the spatial stitching and spectral matching of the single band have been completed. That is to say, the image stitching has been realized for the single band. When each point in the overlapping region of a single band gets a new spectral value, by the same method, all overlapping regions of each band get a new spectral value. The spectral value of the nonoverlapping region still uses the spectral value of the reference image. Therefore, we can get the spectrum after the spectral matching. In order to better illustrate the algorithm process, we show the algorithm flowchart in Algorithm 1.

III. EXPERIMENTS AND ANALYSIS

In this section, we use different remote sensing image datasets to test our proposed method. First, we compare different feature point extraction algorithms, including SIFT [5], SURF [6], and SS-SIFT [7], on hyperspectral image (HSIs) datasets. Then, we compare the feature points matching algorithms, including RANSAC [10], LPM [16], [17], VFC [14], [15], and our mTopKRP, on remote sensing date image set with both rigid and nonrigid transformations. Next, we compare our robust elastic

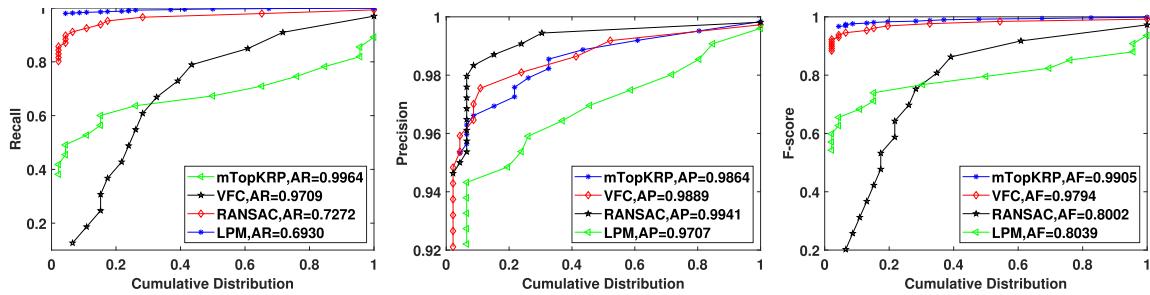


Fig. 4. Quantitative comparisons of RANSAC [10], VFC [14], [15], LPM [16], [17], and our mTopKRP on synthetic-aperture radar (SAR), color infrared aerial photograph (CIAP), and UAV. Recall, Precision, and F-score (left to right) with respect to the cumulative distribution. The average Precision (AP), average Recall (AR), and average F-score (AF) are reported in the legend.

warp with ANAP [21], NISwGSP [22], and ELA [29]. At last, we obtain our panoramic hyperspectral image.

First, we evaluate the feature point extraction algorithm on HSI datasets. This dataset contains 18 HSIs, with the size of 960×1057 . Our data are obtained by GaiaSky-mini push-broom airborne hyperspectral imaging system. Hyperspectral images have a spectral range from 400 to 1000 nm with a total of 176 bands. The shooting height of the image is 300 m. Fig. 2 shows all of our images.

We compare the number of feature points and matching points and time. The results are listed in Table I.

From the aforementioned result, we can find out that SS-SIFT requires too much time, which reduce our efficiency. It shows that SIFT not only can detect more feature points but also require less time. So, we choose SIFT algorithm to detect feature points.

We show the results of SIFT, SURF, and SS-SIFT using the traditional RANSAC algorithm to remove outliers. We find that RANSAC causes the matching points to be concentrated in areas with complex textures (see Fig. 3). It can be concluded that the RANSAC cannot ensure that the matched feature points are evenly distributed in the overlapping region. Besides, some imperceptible wrong matches are still considered correct, which is inevitable.

Then, we compare different false match removal algorithms. The result of false match removal is shown in Fig. 4. The performance values (i.e., Precision, Recall, and F-score) are summarized. The F-score is used for evaluating the matching performance, which is defined as [9]

$$\text{F-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{(\text{Precision} + \text{Recall})} \quad (16)$$

where the Precision is defined as the ratio of the identified correct match number and the preserved match number, and Recall is defined as the ratio of identified correct match number and the correct match number contained in the putative set. Compared to RANSAC, VFC, and LPM algorithm, our mTopKRP algorithm improves F-score. Our method shows its ability to preserve more correct matches, leading to better recalls. We can observe that our method is the best with obvious advantages.

To verify how well the new measuring criterion works, we randomly select in total 46 remote sensing image pairs, such as CIAP, UAV, and SAR. To evaluate our algorithm, we use the following remote sensing image datasets.

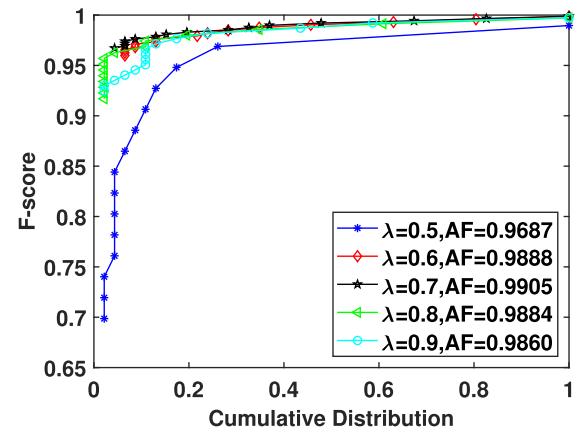


Fig. 5. F-score with respect to the cumulative distribution by using the whole feature set to construct local neighboring structure on 46 remote-sensing image pairs. The best average F-score and its threshold λ are AF = 0.9905 and $\lambda = 0.7$.

- 1) CIAP: This dataset consists of 40 pairs of CIAP, with the size of 700×700 . These images are already orthorectified. Although they only have rigid transformation, the overlap areas are quite small, which is common in image stitching problem.
- 2) UAV: This dataset consists of 35 pairs of color images, which are of resolution 600×337 and captured by a UAV over a piece of farmland. These images may come across projective distortion or nonrigid transformation due to the unstable imaging condition.
- 3) SAR: This dataset contains 34 image pairs corrupted with strong noise. For each image pair, the two images are separately obtained by SARs on a satellite and on an UAV, respectively.

The results are shown in Fig. 5. It can be concluded that our algorithm works best when λ is 0.7 and F-score is 0.9905. Clearly, our method successfully identifies most of the true correspondences, and only a few are wrongly classified. These results prove the generality and robustness of our method.

Next, we compare our result with the result obtained by ANAP [21], NISwGSP [22], and ELA [29] on HSI datasets. The results are shown in Figs. 6 and 7. The former is a gray-scale image after single-band stitching, and the latter is an RGB

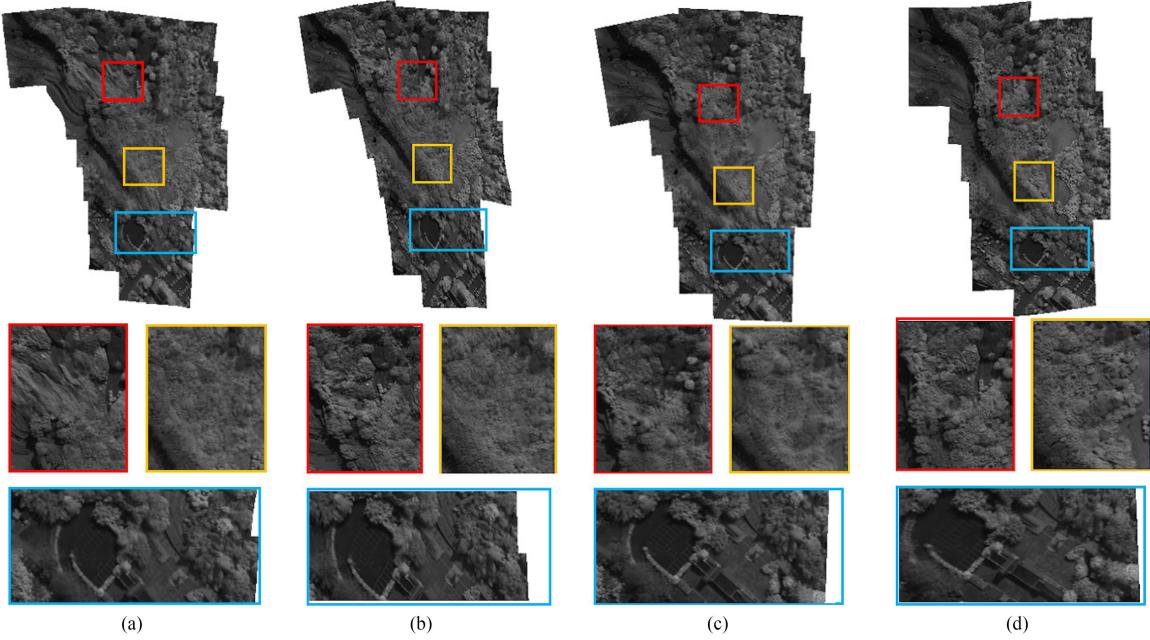


Fig. 6. Qualitative comparison of the stitching performance. The first row is the corresponding stitching results of ANAP [21], NISwGSP [22], ELA [29], and our proposed method of single-band. The rest rows are the three representative areas of each resulting image. (a) Result of ANAP. (b) Result of NISwGSP. (c) Result of ELA. (d) Result of our warp.



Fig. 7. Final hyperspectral stitching result: After the fusion of RGB color image with the 176 bands hyperspectral image at pixel level, the hyperspectral remote sensing image with color display is obtained by our method. We synthesize the pseudocolor image by selecting 92th (700.2 nm), 47th (547.6 nm), and 13th (436.5 nm) bands.

pseudocolor image after all bands fusion. We can see that the images of ANAP, NISwGSP, and ELA after stitching all show obvious distortion, and serious ghosting can be seen after zooming-in. The results show that our method aligns the images more accurately, generating a more natural stitched image with no visible artifacts in the overlapping region and alleviating the perspective distortion problem in nonoverlapping regions.

Traditionally, the evaluation metrics of image mosaic quality mainly include mean value, information entropy, mutual information, SNR, mean grads, etc. In this article, the metrics we used are three functions based on gradient, frequency domain, and space distribution, respectively. EOG is an energy gradient that reflects the change between gray levels. DFT reflects the overall activity of an image space. Variance refers to the discrete degree of gray value of image pixel relative to the mean value. The expressions are as follows:

EOG

$$= \sum_x \sum_y \{[f(x+1, y) - f(x, y)]^2 + [f(x, y+1) - f(x, y)]^2\} \quad (17)$$

$$\text{DFT} = \frac{1}{MN} \sum_{\mu=0} \sum_{\nu=0} \sqrt{\mu^2 + \nu^2} P(\mu, \nu) \quad (18)$$

$$\text{Variance} = \frac{1}{M \times N} \sum_x \sum_y (f(x, y) - \lambda)^2 \quad (19)$$

where $\sqrt{\mu^2 + \nu^2}$ indicates the distance from the pixel to the center pixel, which is used to emphasize the high-frequency components in the spectrum; M and N are the image size; $P(\mu, \nu)$ is the square of the image spectrum; and $\lambda = \frac{1}{M \times N} \sum_x \sum_y f(x, y)$.



Fig. 8. Point A and its surrounding points in I_1 and point A' in I_2 , and point B and its surrounding points in I_1 and point B' in I_2 , which are enlarged to the pixel level.

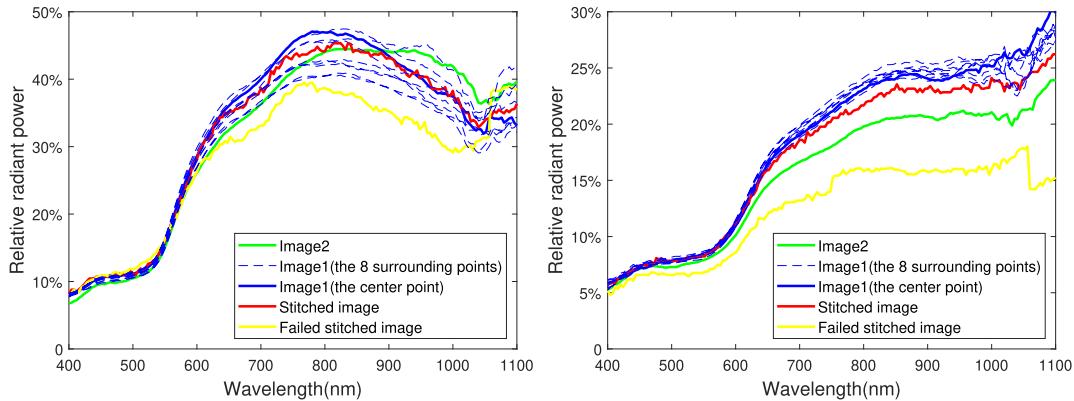


Fig. 9. Spectrum analysis. The left-hand side one uses A and A' representing Image1 and Image2, and the right-hand side one uses B and B' representing Image1 and Image2. The blue and green curves are the spectrum of the two original images before stitching. The blue one is the spectrum of the reference image. The yellow one is the spectrum that recalculates the fusion parameters for each band. The red one is the spectrum after the improved algorithm completes the fitting again.

Table II lists the detailed experiment results. The higher value of EOG, DFT, and Variance means the better quality of the image.

Experiment results show that our warp can achieve good effect. Compared to the ELA, our warp increases the EOG, DFT, and Variance by approximately 7.31%, 39.43%, and 15.77%. Our warp is 7.16%, 32.76%, and 15.23% higher than ANAP, respectively. Our method works well and preserves the naturalness of the image content.

IV. SPECTRUM ANALYSIS

Compared with RGB images, hyperspectral images are special for the spectrum, which is important in the identification or classification of land covers. Therefore, besides the spatial stitching, we also pay attention to the spectral changes of the hyperspectral images after stitching. The following is the spectrum analysis of the overlapping region.

The spectral angle distance (SAD) [33] determines the spectral similarity between two spectra by calculating the angle

between the spectra. This algorithm treats the spectra as vectors in a space with dimensionality equal to the number of bands. By calculating SAD between the test spectrum \mathbf{a}_i and the reference spectrum \mathbf{b}_i , the similarity between them can be estimated

$$\text{SAD} = \arccos \left[\frac{\sum \mathbf{a}_i \mathbf{b}_i}{\sqrt{\sum \mathbf{a}_i^2} \sqrt{\sum \mathbf{b}_i^2}} \right] \quad (20)$$

where i is the number of bands, $\text{SAD} \in [0, 1]$. The smaller the SAD value is, the closer the cosine value is to 1 and the closer the two curves are.

In this experiment, we use HSI datasets. First, the 88th band is taken as an example, and we find the corresponding point pairs of the overlapping areas of the two adjacent images. Then, we record them as point A and A' , B and B' (see Fig. 8). Since the image I_1 is the reference image, the spectrum after fusion should be similar to that of image I_1 . We use SAD to analyze the similarity degree. The difference between the new algorithm and the original one is that the same fusion parameters are used for each band in the linear blending. It can be seen that after the

use of the improved algorithm, the spectrum is close to I_1 . It can be calculated that for the points A and A' , the SAD value of I_1 and I_2 is 0.0894, and the SAD value of the stitched image and I_1 is 0.0308. For B and B' , the SAD value of I_1 and I_2 is 0.0471, and the SAD value of the stitched image and I_1 is 0.0212.

After stitching, the spectrum has obvious sawtooth from the wavelength of 600 nm. We find out the spectrum of I_1 and eight points around the points A and B (see Fig. 9). It can be seen that the spectrum of points A and B in the original image is also quite different from that of the eight surrounding pixels. For A , the SAD values are 0.0562, 0.0617, 0.0273, 0.0411, 0.0301, 0.0175, 0.0388, and 0.0530, respectively. For B , the SAD values are 0.0372, 0.0335, 0.0279, 0.0389, 0.0248, 0.0368, and 0.0331, respectively. It can also be seen from the observation of the spectrum trend that the difference of it between point A and the surrounding points starts from the wavelength of 600 nm and for B , it starts from 700 nm. When using the weighted fusion algorithm, the coordinate value of each pixel in the image is an integer value, but the coordinate point transformed by the geometric transformation model is not necessarily being an integer value. So, we need to get the integer coordinate value of the nearest points of the coordinate point. Thus, the transformed point will be affected by the surrounding points, and its SAD values with the surrounding points are large, which explains the reason why the fitted spectrum is not smooth.

According to the results of spectrum analysis, we can prove that our method guarantees the integrity of the spectral information. It can be drawn from SAD that the spectrum of the stitched image and the reference image is similar, which means the spectrum distortion is very low. Therefore, this method is effective for the stitching of hyperspectral remote sensing images.

V. CONCLUSION

The main contribution of this article is the design of an automatic image stitching algorithm for hyperspectral images, which combines the SIFT algorithm, the mTopKRP algorithm, the robust elastic warps, and the spectral matching algorithm. First, we propose a new mismatch removal method for robust feature matching called mTopKRP, which is based on the stable neighboring topological relationship of feature correspondences. Next, the robust elastic warp is adopted to reduce the distortion of the image. After blending the 18 images for one band, we repeat this process using the same global warp and fusion parameters for the other 175 bands. The mTopKRP algorithm ensures the accuracy of feature point extraction and matching, robust elastic warp avoids some shape distortion, and the spectral matching algorithm makes the spectrum consistent with the original one. Finally, the experimental results on real hyperspectral images demonstrate that our method can obtain high-precision stitching results.

REFERENCES

- [1] A. Plaza *et al.*, “Recent advances in techniques for hyperspectral image processing,” *Remote Sens. Environ.*, vol. 113, pp. S110–S122, 2009.
- [2] J. Jiang, J. Ma, C. Chen, Z. Wang, Z. Cai, and L. Wang, “SuperPCA: A superpixelwise PCA approach for unsupervised feature extraction of hyperspectral imagery,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4581–4593, Aug. 2018.
- [3] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, “Robust feature matching for remote sensing image registration via locally linear transforming,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6469–6481, Dec. 2015.
- [4] X. Jiang, J. Ma, J. Jiang, and X. Guo, “Robust feature matching using spatial clustering with heavy outliers,” *IEEE Trans. Image Process.*, vol. 29, pp. 736–746, 2019.
- [5] D. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, pp. 91–110, Nov. 2004.
- [6] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “SURF: Speeded up robust features,” *Comput. Vis. Image Understanding*, vol. 110, pp. 346–359, Jan. 2008.
- [7] A.-K. S. Lateef, J. Zhou, A. Zia, and W. C. Liew, “Spectral-spatial scale invariant feature transform for hyperspectral images,” *IEEE Trans. Image Process.*, vol. 27, no. 2, pp. 837–850, Feb. 2018.
- [8] S. M. M. Kahaki, M. J. Nordin, and A. H. Ashtari, “Contour-based corner detection and classification by using mean projection transform,” *Sensors*, vol. 14, pp. 4126–4143, 2014.
- [9] S. M. M. Kahaki, M. J. Nordin, A. H. Ashtari, and S. J. Zahra, “Deformation invariant image matching based on dissimilarity of spatial features,” *Neurocomputing*, vol. 175, pp. 1009–1018, 2016.
- [10] M. A. Fischler and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [11] S. M. M. Kahaki, S.-L. Wang, and A. Stepaniants, “Accurate registration of in vivo time-lapse images,” *Med. Imag.: Image Process.*, vol. 10949, pp. 345–350, 2019.
- [12] P. H. Torr and A. Zisserman, “MLESAC: A new robust estimator with application to estimating image geometry,” *Comput. Vis. Understanding*, vol. 78, no. 1, pp. 138–156, 2000.
- [13] O. Chum and J. Matas, “Matching with PROSAC—Progressive sample consensus,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 220–226.
- [14] J. Ma, J. Zhao, J. Tian, A. L. Yuille, and Z. Tu, “Robust point matching via vector field consensus,” *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1706–1721, Apr. 2014.
- [15] J. Ma, J. Zhao, J. Tian, X. Bai, and Z. Tu, “Regularized vector field learning with sparse approximation for mismatch removal,” *Pattern Recognit.*, vol. 46, no. 12, pp. 3519–3532, 2013.
- [16] J. Ma, J. Zhao, J. Jiang, H. Zhou, and X. Guo, “Locality preserving matching,” *Int. J. Comput. Vis.*, vol. 127, no. 5, pp. 512–531, 2019.
- [17] J. Ma, J. Jiang, H. Zhou, J. Zhao, and X. Guo, “Guided locality preserving feature matching for remote sensing image registration,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4435–4447, Aug. 2018.
- [18] M. Brown and D. G. Lowe, “Automatic panoramic image stitching using invariant features,” *Int. J. Comput. Vis.*, vol. 74, no. 1, pp. 59–73, 2007.
- [19] J. Zaragoza, T.-J. Chin, Q.-H. Tran, M. S. Brown, and D. Suter, “As-projective-as-possible image stitching with moving DLT,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 7, pp. 1285–1298, Jul. 2014.
- [20] C.-H. Chang, Y. Sato, and Y.-Y. Chuang, “Shape-preserving half-projective warps for image stitching,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 3254–3261.
- [21] C.-C. Lin, S. U. Pankanti, K. N. Ramamurthy, and A. Y. Aravkin, “Adaptive as-natural-as-possible image stitching,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1155–1163.
- [22] Y.-S. Chen and Y.-Y. Chuang, “Natural image stitching with the global similarity prior,” in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 186–201.
- [23] C. Yuan, M. Zhou, L. Sun, S. Qiu, and Q. Li, “A multi-dimensional hyperspectral image mosaic method and its acquisition system,” in *Proc. Int. Congr. Image Signal Process. BioMed. Eng. Inf.*, 2018, pp. 1–6.
- [24] J. Fang, X. Wang, T. Zhu, X. Liu, X. Zhang, and D. Zhao, “A novel mosaic method for UAV-based hyperspectral images,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 9220–9223.
- [25] X. Jiang, J. Jiang, A. Fan, Z. Wang, and J. Ma, “Multiscale locality and rank preservation for robust feature matching of remote sensing images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6462–6472, Sep. 2019.
- [26] J. Li, P. Jiang, S. Song, H. Xia, and M. Jiang, “As-aligned-as-possible image stitching based on deviation-corrected warping with global similarity constraints,” *IEEE Access*, vol. 7, pp. 156603–156611, 2019.
- [27] J. Ma, J. Wu, J. Zhao, J. Jiang, H. Zhou, and Q. Z. Sheng, “Nonrigid point set registration with robust transformation learning under manifold regularization,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3584–3597, Dec. 2019.
- [28] J. Ma, J. Zhao, Y. Ma, and J. Tian, “Non-rigid visible and infrared face registration via regularized Gaussian fields criterion,” *Pattern Recognit.*, vol. 48, no. 3, pp. 772–784, 2015.

- [29] J. Li, Z. Wang, S. Lai, Y. Zhai, and M. Zhang, "Parallax-tolerant image stitching based on robust elastic warping," *IEEE Trans. Multimedia*, vol. 20, no. 7, pp. 1672–1687, Jul. 2018.
- [30] K. Sun and Y. Wan, "Image stitching with weighted elastic registration," in *Proc. 10th Int. Conf. Commun., Circuits Syst.*, Dec. 2018, pp. 403–407.
- [31] F. Lei and W. Wang, "A fast method for image mosaic based on surf," in *Proc. 9th IEEE Conf. Ind. Electron. Appl.*, 2014, pp. 79–82.
- [32] J. Gao, S. Kim, and M. Brown, "Constructing image panoramas using dual-homography warping," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2011, pp. 49–56.
- [33] F. Kruse *et al.*, "The spectral image processing system (SIPS) interactive visualization and analysis of imaging spectrometer data," *Remote Sens. Environ.*, vol. 44, no. 2, pp. 145–163, 1993.



Yujie Zhang is currently working toward the B.E. degree with the Electronic Information School, Wuhan University, Wuhan, China.

She is currently a Research Assistant with the Multi-Spectral Vision Processing Laboratory, Wuhan University. Her current research interests include hyperspectral imagery and image processing.



Zhiying Wan received the B.E. degree from the Electronic Information School, Wuhan University, Wuhan, China, in 2020.

She was a Research Assistant with the Multi-Spectral Vision Processing Laboratory, Wuhan University. Her current research interests include hyperspectral imagery and image processing.



Xingyu Jiang received the B.E. degree from the Department of Mechanical and Electronic Engineering, Huazhong Agricultural University, Wuhan, China, in 2017, and the M.S. degree in 2019 from the Electronic Information School, Wuhan University, Wuhan, China, where he is currently working toward the Ph.D. degree.

His research interests include computer vision, machine learning, and pattern recognition.



Xiaoguang Mei received the B.S. degree in communication engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2007, the M.S. degree in communications and information systems from Huazhong Normal University, Wuhan, China, in 2011, and the Ph.D. degree in circuits and systems from the HUST, in 2016.

From 2010 to 2012, he was a Software Engineer with the 722 Research Institute, China Shipbuilding Industry Corporation, Wuhan, China. From 2016 to 2019, he was a Postdoctoral Fellow with the Electronic Information School, Wuhan University, Wuhan, China, where he is currently an Assistant Professor. His research interests include hyperspectral imagery, machine learning, and pattern recognition.