Article

Large-Scale Stitching of Hyperspectral Remote Sensing Images Obtained from Spectral Scanning Spectrometers Mounted on Unmanned Aerial Vehicles

Hong Liu [1,2], Bingliang Hu [1,*], Xingsong Hou [2], Tao Yu [1], Zhoufeng Zhang [1], Xiao Liu [1], Xueji Wang [1] and Zhengxuan Tan [3]

1   Key Laboratory of Spectral Imaging Technology, Xi'an Institute of Optics Precision Mechanics of Chinese Academy of Sciences, Xi'an 710119, China; liuhong@opt.ac.cn (H.L.); yutao@opt.ac.cn (T.Y.)
2   School of Electronic and Information Engineering, Xi'an Jiao Tong University, Xi'an 710049, China; houxs@xjtu.edu.cn
3   Department of Computer Sciences, University of Miami, Miami, FL 33136, USA
*   Correspondence: hbl@opt.ac.cn

Abstract: To achieve large-scale stitching of the hyperspectral remote sensing images obtained by unmanned aerial vehicles (UAVs) equipped with an acousto-optic tunable filter spectrometer, this study proposes a method based on a feature fusion strategy and a seam-finding strategy using hyperspectral image classification. In the feature extraction stage, SuperPoint deep features from images in different spectral segments of the data cube were extracted and fused. The feature depth matcher, LightGlue, was employed for feature matching. During the data cube fusion stage, unsupervised K-means spectral classification was performed separately on the two hyperspectral data cubes. Subsequently, grayscale transformations were applied to the classified images. A dynamic programming method, based on a grayscale loss function, was then used to identify seams in the transformed images. Finally, the identified splicing seam was applied across all bands to produce a unified hyperspectral data cube. The proposed method was applied to hyperspectral data cubes acquired at specific waypoints by UAVs using an acousto-optic tunable filter spectral imager. Experimental results demonstrated that the proposed method outperformed both single-spectral-segment feature extraction methods and stitching methods that rely on seam identification from a single spectral segment. The improvement was evident in both the spatial and spectral dimensions.

Keywords: hyperspectral image (HSI) stitching; deep feature extraction; feature matching; acousto-optic tunable filter; unmanned aerial vehicle (UAV)

1. Introduction

Drone technology is a commonly used industrial technology today, and hyperspectral imaging is an effective combination of spectral and video camera technologies. Unmanned aerial vehicle (UAV) hyperspectral remote sensing can achieve target scene acquisitions while maintaining spectral resolution ability, making it an effective means for Earth observations. Compared with other hyperspectral remote sensing sensor platforms, UAV hyperspectral remote sensing technology has the advantages of simple operation, adjustable flight path and altitude, and convenient data acquisition. It has been extensively used in geological and mineral research [1,2], forest resource investigation [3], environmental monitoring [4,5], agricultural assessment [6], and other fields.

Why do we need to perform the large-scale concatenation of hyperspectral data cubes? Although drone hyperspectral remote sensing technology has broad application prospects due to its low cost, high spectral resolution, and great flexibility [7–9], the limitations of drone flight altitude and short single flight operation time result in a relatively small range of a single field of view and single flight scenery [10]. To obtain a large panoramic-ranged hyperspectral image, it is necessary to perform large-scale stitching of the hyperspectral data cubes [11].

There are various ways to achieve hyperspectral image stitching. Traditional image stitching typically involves the following steps: feature extraction, matching and alignment, fusion, and optimization [12]. The common steps for achieving hyperspectral data cube stitching at present include the following [13]: (a) identifying the reference spectral segment, (b) feature extraction and matching, (c) image alignment, and (d) its application to all spectral segments. Feature extraction and matching are used to calculate homography. By obtaining a sufficient number of correct matching features, a homography estimation matrix is obtained, and the homography estimation is distorted to obtain the registered image. The aligned images also need to be stitched together. Due to the inevitable impact of projection distortion on homography regularization, many methods have been proposed to eliminate ghosting in stitched images. The image alignment methods can be divided into non-seam and seam stitching based on whether seam prediction is performed or not.

Feature extraction methods include SIFT and SURF [14]. To ensure the accuracy of the matching results, it is necessary to detect and remove outliers in feature points. Random sampling consensus [15] is a classic method to remove false matches, but it cannot be applied to specific nonrigid and complex situations. VFC [16], LPM [17], and other methods are suitable for nonrigid motion scenes. In addition, Zhang Yujie et al. [18] extracted SIFT feature points during the stitching of hyperspectral images carried by drones, matched the feature points, and removed outliers using the mTopKRP method. Numerous experiments have demonstrated the effectiveness of the proposed method, which utilizes robust feature matching and elastic warping for the automatic stitching of hyperspectral images. Zhang et al. [19] used the deep-learning feature SuperPoint and implemented feature point matching and error removal through the LAF [20] algorithm to enhance the reliability of feature correspondence. Zhiying Jiang et al. [21] utilized the advantages of graph convolutional networks in modeling feature relationships and proposed a multispectral image stitching method based on spatial graph inference. This method can effectively extract features from multispectral images from different viewpoints.

Fusion ghosting from non-seam stitching is caused by the fact that the homography matrix cannot align the two images perfectly, and ghosting is eliminated by aligning the target with the reference image as much as possible. The APAP method [22] places a grid in the image and estimates a local homography model for each grid. Owing to its excellent performance, this technology has been extensively applied in image alignment. To achieve better alignment, the robust ELA method [23] combines a grid-based model with a direct deformation strategy. To maintain good performance in low-texture environments, Li Nan et al. [24] developed a dual-feature distortion model for image alignment, which utilizes sparse feature matching and line correspondence simultaneously. The aforementioned methods divide the image into different regions and calculate the homography matrix for each region. By applying spatial distortions on certain areas, overlapping regions are well aligned and ghosting is reduced considerably. However, Yang Lichun et al. [25] found that non-seam stitching methods performed well in small disparity images, but were difficult to use for handling large disparity images, which could lead to poor stitching results.

Seam stitching hides artifacts by studying the optimal seam for stitching distorted images and improves ghosting issues by optimizing costs associated with seams. Senmao

Cheng et al. [26] proposed a deep-learning-based seam prediction method. Qi Jia et al. [27] proposed a seam-matching strategy that utilizes line-point consistency measurements. Liao et al. [28] proposed a new iterative seam estimation method that uses a mixed-quality evaluation method to evaluate pixels along the seam. Li et al. [29] proposed a quaternion rank 1 alignment model that simultaneously learns the optimal seam line and local alignment. Nie et al. [30] proposed a seam stitching method based on deep learning, but there were various issues, such as unclear boundaries and discontinuous generated masks. Li et al. [13] proposed using the edge-enhancement energy function for orthorectified image stitching to detect optimal seams.

There are two difficulties in achieving hyperspectral data stitching. The first difficulty is the identification of ways to select reference spectral segments for the feature extraction of two data cubes, the identification of the types of features that need to be extracted, and the determination of ways to pair features. The second difficulty is associated with the fact that hyperspectral data contain a large number of spectral segments. Accordingly, what are the ways to handle the overlapping parts of two data cubes? The currently popular method is to find a seam in the overlapping area based on the reference spectral segment. That is, the determination of ways to identify a seam that has advantages in both spatial and spectral dimensions of the concatenated data is another difficulty in hyperspectral data stitching.

Regarding the first difficulty, the selection of reference bands is the first step in hyperspectral image stitching, and the quality and effectiveness of the stitching are directly related to the selection of the reference bands. The current research provides four methods for selecting benchmark spectral bands:

Zhang et al. [18] extracted images from the 88th band in the middle of 176 bands as the feature extraction images without considering images from other spectral bands. This selection method is simple, but selecting feature extraction spectral bands is not representative. Zhang et al. [19] found the spectral segment with the highest signal-to-noise ratio. Although this method found the spectral segment with the highest signal-to-noise ratio, it may not necessarily be the spectral segment with the most features. Mo et al. [31] divided the entire wavelength range into three segments, simply added the images of different spectral segments, and then multiplied them by the coefficients to the weight to synthesize a single image. This image was used to extract features, and the resulting image weakened the spectral information. Li et al. [32] adopted a baseline band selection method based on prior knowledge and PCA, using the product of mean and standard deviation as a threshold to select spectral segments greater than the threshold. This method considers the integrity of each spectral segment, but whether the selected first principal component can represent all spectral segments still needs to be verified.

Regarding the second challenge, another important step in stitching images is to find the seam line within the overlapping area of the image, which is the line connecting the most similar pixels in the overlapping area. After determining the seam line in the overlapping area, only the image parts on one side and the other side of the line are selected instead of simply merging the two images of the overlapping area. The purpose of doing so is to avoid blurring and introducing artifacts in the image. Previous researchers have proposed the following methods for obtaining and using splicing seams:

Peng et al. [33,34] used a fast and robust seam estimation method to determine the seams in the single bands of hyperspectral images. This method applies to red–green–blue (RGB) images, which can be modified into single-band grayscale images for hyperspectral imaging. To ensure the integrity of spatial and spectral information of the hundreds of bands in hyperspectral images, the structural similarity index (SSIM) is applied to select the optimal one among all band candidate stitching lines, and the selected optimal stitching

line is applied to the remaining bands for stitching. Mo et al. [31] also used a method to identify the optimal splicing seam [35], which relied on the input single-spectral segment. They used the graph-cutting method to obtain a splicing seam and applied it to other spectral segments. This method has been proven to be an effective method for parallax image stitching.

In this study, we aim to propose a novel method for large-scale stitching of the hyperspectral data cube. The main contributions of this study are as follows:

First, we introduce a new strategy for multi-feature fusion, which considers the features of all spectral segments as the overall features of the data cube. These fused features comprehensively reflect the characteristics of each spectral segment of the data cube, overcoming the shortcomings of traditional methods that use single-spectral images to replace features in data cubes with insufficient feature amounts.

Second, for the data cube fusion stage, we propose a seam-finding strategy based on hyperspectral image classification. This method addresses the drawback of using only a single spectral segment to search for splicing seams without fully utilizing the spectral information, thus reducing inaccuracies in splicing seams.

The remainder of this paper is organized as follows: Section 2 outlines the related work. Section 3 explains the proposed method for hyperspectral data cube stitching. Section 4 presents the experimental results and Section 5 discusses them. Finally, Section 6 concludes this study.

2. Related Work

This section provides an overview of the proposed method. First, the airborne acousto-optic tunable filter (AOTF) spectrometer, which utilizes a zoom lens, is introduced. Subsequently, the waypoint collection process is described. Finally, the route and waypoint collection mode is explained.

2.1. Unmanned Aerial Hyperspectral Imaging System Based on AOTF Spectrometer

With the rapid advancements in materials science and the increasing sophistication of optical device manufacturing, spectrometers have evolved from traditional prism- or grating-based designs to modern devices featuring higher spectral resolution, compactness, integration, and intelligence [36]. The AOTF, an all-solid-state filtering and polarization modulation device, offers several advantages, including fast electronic tuning, the absence of mechanical moving parts, a compact structure, high-diffraction efficiency, and a broad tuning range. These characteristics have facilitated its application in the development of spectral imaging systems [37,38].

The UAV-mounted hyperspectral imaging system, based on an AOTF spectrometer (China Electronics Technology Group 26th Research Institute, Chongqing, China) (Figure 1a), comprised a lightweight UAV, an AOTF imaging spectrometer with an electrically controlled zoom lens (Figure 1b), an AOTF driver, a mini personal computer (MINI-PC), and a battery.

The AOTF spectrometer was based on an electric zoom lens with a spectral range of 400–1000 nm; a spectral resolution of 8.2 nm (at the 450 nm band), 8.5 nm (at the 650 nm band), 10 nm (at the 850 nm band); and is variable at each center wavelength. The acquisition software was set to 121 bands, ranging from 400 to 1000 nm, with a 5 nm step size increase. The horizontal resolution was 2048, vertical resolution was 680, and the quantization bit count was 16 bits.

Figure 1. Composition of the unmanned aerial vehicle (UAV)-mounted hyperspectral imaging system based on the airborne acousto-optic tunable filter (AOTF) spectrometer. (**a**) Components of the imaging system included the AOTF spectrometer, AOTF driver, MINI-PC, and battery. (**b**) AOTF spectrometer airborne imaging system featuring a zoom lens.

The drones were equipped with spectrometers, AOTF drivers, batteries, and other components. The drone platform adopted the DJI WIND4 customized flight platform, comprising a motor wheelbase of 1050 mm, a total weight (excluding batteries) of 7.3 kg, a maximum takeoff weight of 24.5 kg, a maximum flight speed of 14 m/s, and a hover time (single battery) of 28 min. However, considering the duration of the ascent and descent phases and the safety buffer, the planned flight duration typically did not exceed 20 min.

The acquisition and control computer comprised the following: implemented AOTF spectrometer data acquisition, AOTF driver control, and zoom lens control. The control computer adopted Intel's NUC mini-host and i5-7260U processor, and had a main frequency in the range of 2.2–3.2 GHz, a 4 MB high-speed cache, 15 W thermal design power consumption, 32 GB memory, a size of 117 $\times$ 112 $\times$ 52 mm, and a 200 G solid-state drive storage hard drive that was capable of meeting the storage needs of 500 waypoint data cubes.

The core optical path structure of the AOTF spectrometer with an electrically controlled zoom lens is illustrated in Figure 2. It included the following components: (1) an electrically controlled zoom lens, (2) a front objective lens, (3) an aperture stop, (4) a collimating lens, (5) a linear polarizer, (6) an AOTF module comprising a TeO$_2$ crystal and a piezoelectric transducer, (7) a second linear polarizer, (8) a secondary imaging lens,

(9) a complementary metal-oxide semiconductor (CMOS) detector, (10) a MINI-PC for control and data acquisition, and (11) a radiofrequency (RF) driver.
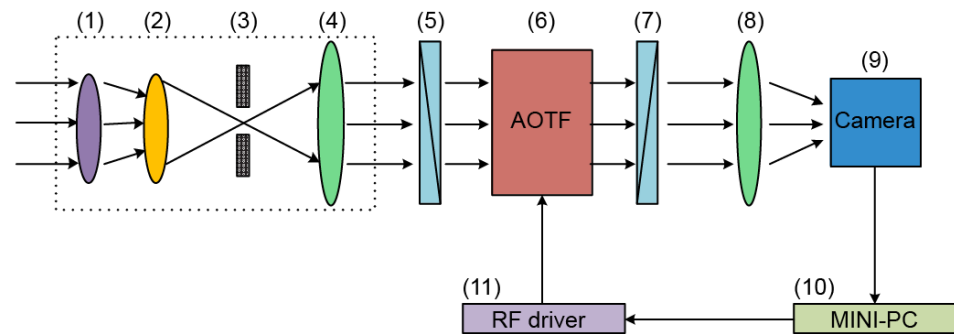
Figure 2. Core optical path structure of the AOTF spectrometer detailing the arrangement of the key components.

The system functioned as follows: the incident light beam passed through the electrically controlled zoom lens (1), front objective lens (2), aperture stop (3), and collimating lens (4) before reaching the surface of the linear polarizer (5). The polarized light was then reflected at a vertically incident on the AOTF module (6), where it interacted with ultrasonic waves to produce a diffracted beam. After passing through the second linear polarizer (7), the beam was focused by the secondary imaging lens (8) onto the CMOS detector (9), where it was collected and processed by the MINI-PC (10). Notably, the polarization direction of the first linear polarizer (5) was parallel to the acousto-optic interaction plane of the AOTF and perpendicular to the polarization direction of the second linear polarizer (7). This configuration ensured that the zero-order transmitted light was filtered out.

2.2. Remote Sensing Waypoint Flight Acquisition Control

The flight route and waypoints were planned based on the remote sensing data collection area using DJI's ground flight control assistance software. The selection of flight routes and waypoints must account for the UAV's endurance time. Waypoints were set by considering the ground field of view, calculated using the focal length and aperture stop of the spectrometer's imaging objective. Overlapping ground field-of-view ranges between the consecutive waypoints were necessary to ensure seamless stitching of large-scale spectral data. After defining the route and waypoints, specific actions needed to complete the spectrometer data cube acquisition were assigned to each waypoint, including the triggering of the RGB camera exposure, starting the imaging spectrometer scan, and the specification of the drone's hover time. The flight plan, including the route and waypoint information, was exported as a keyhole markup language file using the ground flight control assistance software and imported into the drone flight control software for execution. Figure 3 summarizes the process followed to complete the UAV hyperspectral remote sensing data acquisition.

The detailed process was as follows:

(1) Before the remote sensing flight test, system connections were checked. The system was powered on, and the control computer (MINI-PC) was started. Remote sensing data acquisition control software was launched, linking the AOTF-based hyperspectral imager and RF controller, followed by a ground photography self-test.

(2) Parameters such as the integration time and gain settings for the spectral camera were configured based on weather conditions. The spectral range and number of spectral segments for the AOTF spectrometer were also set.
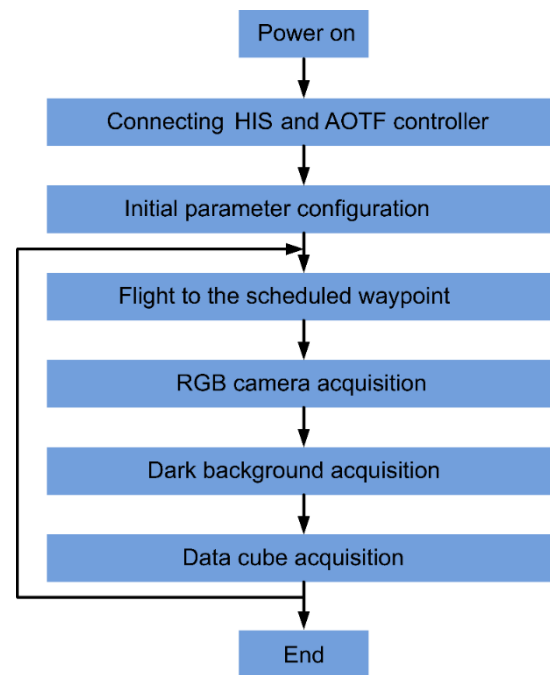
Figure 3. Flowchart of the UAV remote sensing data acquisition process.

(3)   Preset route and waypoint data, including latitude, longitude, altitude, yaw angle, and hover time, were imported into the drone flight control software. The drone autonomously executed the flight plan.

(4)   Upon reaching a waypoint, the UAV triggered the RGB panchromatic camera to capture a wide-field image. Simultaneously, the spectrometer acquisition control program received the trigger command.

(5)   The spectrometer acquisition control program directed the AOTF driver to disable the RF drive signal to acquire a dark background image.

(6)   The spectrometer program performed the high-speed acquisition of wavelength-specific data based on the configured parameters. A pre-established distortion model was applied during acquisition for real-time correction, and the hyperspectral data cube was stored in a designated format.

(7)   The system checked if the final waypoint had been reached. If not, it proceeded to the next waypoint and repeated steps 4–7. If the route was complete, the flight ended, and the UAV returned to the ground station.

Unlike push-broom scanning spectrometers, the AOTF-based imaging spectrometer captures spectral data by hovering at each waypoint, completing the spectral scan for a single data cube. To accommodate this mode of operation, the DJI drone flight controller was configured to synchronize its operation with external trigger signals, and a camera waypoint trigger acquisition program was developed. This program enabled rapid switching and scanning of spectral segments. In route planning, adherence to aviation photography standards and ensuring smooth transitions between the routes were critical for successful image stitching.

2.3. Large Field-of-View Stitching Based on Waypoint Remote Sensing Images

Using the spectral scanning capabilities of the AOTF spectrometer, remote sensing spectral data over a large spatial range were acquired by collecting individual data cubes at multiple waypoints. At each waypoint, a data cube was captured directly over the field of view. The drone sequentially moved to adjacent fields of view to acquire additional data cubes, which were subsequently concatenated to form a larger field of view. Traditional

large-field stitching involved several steps: first, feature point extraction was performed on each image; next, feature point matching and image registration were performed; subsequently, the images were copied to specific positions in a target image; and finally, the overlapping boundaries were processed to achieve seamless integration. For hyperspectral data cubes, stitching was conducted on a per-spectral-segment basis, concatenating the data cubes sequentially. As illustrated in Figure 4, the UAV collected a hyperspectral data cube at each waypoint along its flight path, with each cube representing a distinct field of view.
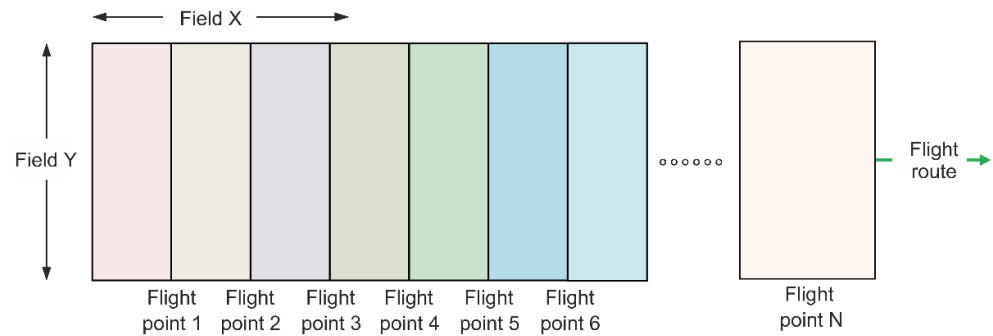
Figure 4. Field-of-view splicing.

When the drone operated at a flight altitude of 100 m and the field-of-view aperture was 10 mm $\times$ 3.2 mm, the focal length of the front objective lens was set to 12 mm to ensure high-spatial resolution. Under these conditions, the ground field of view measured approximately 83 m $\times$ 26 m. To meet the requirements for the field-of-view stitching and following relevant aerial photography technical specifications [39–41], two adjacent fields of view must have a considerable overlap. Typically, the heading overlap rate ranges from 53% to 65%, whereas the lateral overlap rate falls between 15% and 40%. In practical applications, especially for large fields of view with fewer surface features, both heading and lateral overlap rates should be increased to ensure high-quality image stitching. Consequently, to maintain adequate overlap, the distance set during route planning between two consecutive waypoints did not exceed 15 m.

3. Proposed Method

Current methods for data cube concatenation rely solely on features from a single spectral segment, neglecting the information from other spectral segments. Consequently, this study proposes a concatenation strategy based on multi-feature fusion and a stitching seam-search strategy based on spectral image classification. This section outlines the Super-Point feature extraction and multi-feature fusion, LightGlue feature matching, K-means spectral image classification, and the dynamic programming method for finding seams.

Figure 5 shows a schematic of the proposed hyperspectral data cube concatenation method. Initially, SuperPoint depth features were extracted for each spectral segment in Data Cube 1 and Data Cube 2. Subsequently, features from different spectral segments in the two data cubes were fused separately, employing feature addition rather than pixel addition. Subsequently, by using the LightGlue deep-learning feature-matching method, features in the two cubes were paired, incorrect pairings were eliminated, and the transformation matrix was computed. This matrix was then utilized to transform each spectral segment in the two data cubes. Simultaneously, the K-means spectral classification algorithm was employed to classify hyperspectral images in both data cubes, yielding classified images. The grayscale of the classified image was transformed, the transformation matrix was applied to the transformed classified image, and the dynamic programming method—based on a grayscale loss function—was utilized to compute the stitching seam

of the classified image. Finally, the stitching seam of the classified image was applied to all spectral segments of the data cube, achieving the concatenation of the two data cubes.
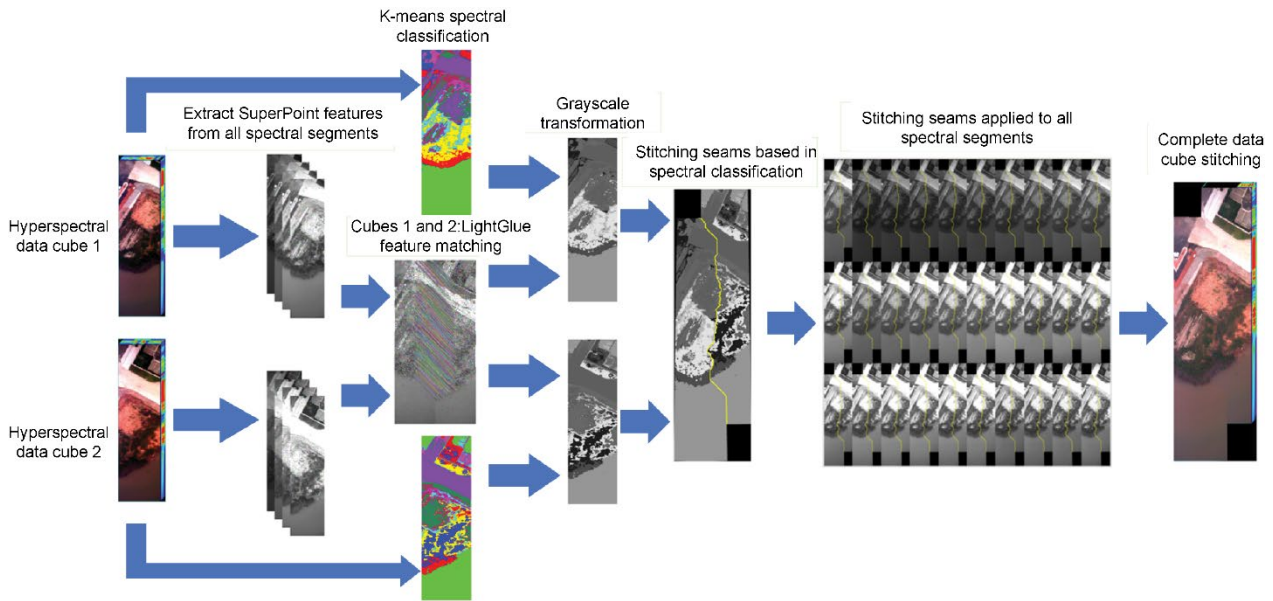

Figure 5. Proposed hyperspectral data cube concatenation method.

3.1. SuperPoint Feature Extraction and Multi-Feature Fusion

SuperPoint is a fully convolutional neural network (CNN) architecture [42] designed to predict feature points and descriptors in a single forward pass through CNNs and is capable of handling inputs of full-size images. The encoder within this model synchronized, parallelized, and shared information. The output image was convolved by the encoder to obtain a feature map with reduced dimensions. The architecture comprised two decoders: one for extracting the key detectors of feature points and another for predicting descriptors. The SuperPoint network could perform two tasks simultaneously and share information, thus addressing the lack of sharing and representation between the points of interest and descriptors. The SuperPoint network architecture is illustrated in Figure 6.
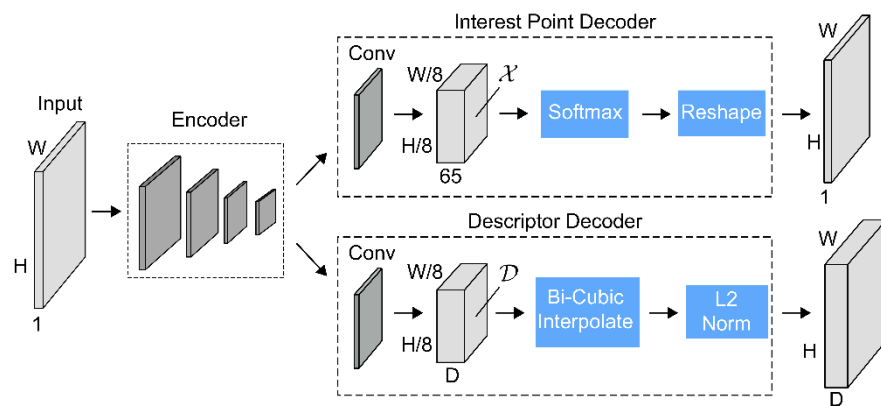

Figure 6. Architecture of the SuperPoint network [42].

SuperPoint, based on deep-neural networks, trains a feature point detector and descriptor capable of self-monitoring learning. It effectively extracts key points in various complex environments and generates feature descriptors with strong robustness for accurate image matching, crucial for alignment in image stitching.

The higher the number of features identified, the greater the number of paired features in the two data cubes, resulting in more matched feature pairs and better concatenation.

SuperPoint features were computed to fully utilize the spectral feature information of different spectral bands. The features of different spectral segments were stored in vector form, which were then aggregated to obtain the features of the cube. The formula describing the multi-feature fusion for the cube is given by

$$Vector < KeyPoint > Cube = \sum_{i}^{n} Vector < KeyPoint > Image(i) \qquad (1)$$

where $Vector < KeyPoint > Cube$ represents the set of all spectral features in the data cube, $Vector < KeyPoint > Image(i)$ represents all features contained in the *i*th spectral image, and *n* represents the number of spectral segments contained in the data cube.

3.2. LightGlue Feature Matching

LightGlue is a deep-learning-based local feature-matching method [43], derived from SuperGlue [44], designed to predict partially matching relationships between local feature sets extracted from images A and B. It consisted of L identical feature processing layers stacked together to process the feature sets of images A and B. Each layer comprised self-attention and self-cross-attention units, with a classifier module introduced at the end of each layer to determine whether to halt the inference and prevent unnecessary calculations. The network structure of LightGlue is depicted in Figure 7.
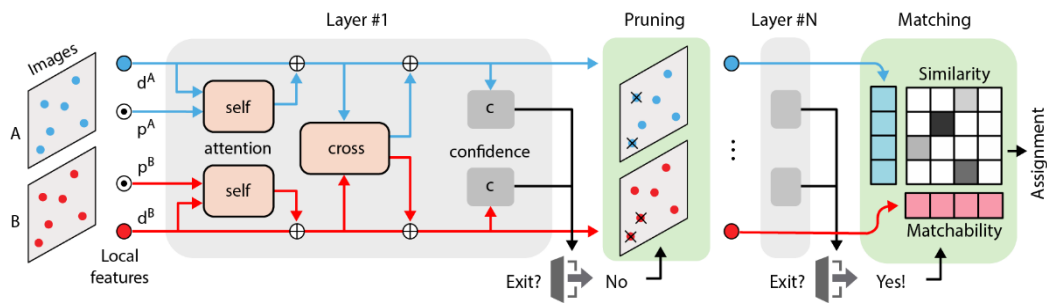

Figure 7. The network structure of LightGlue [43].

LightGlue served as a versatile image registration library primarily focused on optimizing image geometric transformations and stitching. Leveraging the feature point information provided by SuperPoint, it could compute the optimal image fusion strategy, facilitating smooth transitions and eliminating discontinuities in overlapping areas.

3.3. K-Means Spectral Image Classification

The K-means clustering algorithm is extensively employed in spectral image classification. Its fundamental principle involves clustering samples into *K* categories based on similarity, utilizing an initial *K* value and *K* initial cluster centers. The algorithm aims to minimize the Euclidean distance between each sample and its cluster center until it is less than that of other cluster centers. Subsequently, the cluster centers are updated based on distance calculations. This allocation and updating process iterates until the distance between all samples in each cluster and the center point minimally changes, ensuring cluster stability.

Assuming that the hyperspectral image data comprise *N* spectral segments, with each spectral segment image consisting of *M* pixels, the spectral curve of the ith pixel can be

represented by a vector $X_i = \{x_{i,1}, x_{i,2}, \ldots, x_{i,N}\}$, where $M \geq i \geq 1$. The Euclidean distance between the ith and jth samples can be expressed as

$$d(i,j) = \sqrt{\sum_{n=1}^{N} (x_{i,n} - x_{j,n})^2} \qquad (2)$$

The spectral differences for each spectral segment were calculated using equal weights and then summed to determine the overall differences between the spectral segments of the two vectors. This difference, known as similarity, served as a criterion to determine whether sample points belonged to the same class. Smaller distances indicated greater similarity.

The criterion function used to determine the number of iterations in clustering is typically represented by an error function. When using Euclidean distance as the similarity measure and assuming that the clustering results were grouped into K classes, the criterion function can be expressed as

$$J(X,V) = \sum_{k=1}^{K} \sum_{X_i \in V_k} d^2(X_i, V_k) \qquad (3)$$

where V_k represents the center of the kth class.

Assuming C_k is the set of all sample data for the kth class, then

$$C_k = \left\{ X_i \in S \middle| k = argmind^2(X_i, V_k) \right\} \qquad (4)$$

Therefore, the cluster center is

$$V_k = \frac{\sum_{x_i \in C_k} X_i}{|C_k|} \qquad (5)$$

3.4. Dynamic Programming Method for Finding Joint Seams

When there was an overlap between images I_1 and I_2, it was essential to determine the optimal seam between them. Initially, we defined the energy function of the overlap as

$$e = \|I_1, I_2\|_2 \qquad (6)$$

where I_1 and I_2 represent the overlapping parts of the two images, and $\| \ \|_2$ represents the two paradigms.

Typically, there are three constraints to the seam line. Firstly, if the overlapping area is wider than it is tall, the seam will run horizontally; conversely, if the overlap area has a higher height than width, the seam will run vertically, ensuring a definite length for the seam line. Secondly, if the seam is horizontal, it cannot have an absolute vertical seam line, and vice versa. Thirdly, if the overlapping area forms a rectangle, the seam line starts from one side of the rectangle and ends on the side parallel to the starting side.

Assuming that the width of the overlapping area was less than the height, the seam line was vertical. The specific implementation steps were as follows:

1. Each pixel in the first row and column of the energy function corresponded to a stitching line, with its energy value initialized to its current energy value.
2. Starting from the second row, one of the best path nodes for each point was selected in the same row. The selection method involved comparing the energy values of the three adjacent points in the row opposite the current point, recording the column

corresponding to the minimum value, and adding this minimum value to the energy value corresponding to the current point to update the suture's energy value

$$E_{i,j} = e_{i,j} + min(E_{i-1,j-1}, E_{i-1,j}, E_{i-1,j+1}) \tag{7}$$

where $e_{i,j}$ represents the current energy value of the suture, and the three terms in the min function represent the energy values of the current pixel's upper left, upper, and upper right neighbors, respectively.

3. If the current point of the suture line was the last in the figure, the method proceeded to step 4. Otherwise, the method returned to step 2 and continued the expansion.

4. The minimum value in the last row indicated that the end of the minimum vertical path had been reached, allowing us to trace back and obtain the optimal path, which formed the seam line.

Similarly, if the width of the overlapping area was greater than the height, a similar calculation process was performed. The dynamic programming algorithm was well-suited for solving Equation (6).

4. Experimental Results

Two hyperspectral data cubes were captured using a drone equipped with an AOTF spectrometer [37,45]. Owing to the system design, the spectral response efficiency was low at 400–450 nm and 800–1000 nm. To validate the proposed method, we focused on the spectral range of 450–800 nm, with spectra increasing in steps of 5 nm. In total, 71 spectral segments were chosen for the experimental data. Consequently, the size of both data cubes was 530 pixels × 2000 pixels × 71 bands. Three datasets were utilized: the Lakeside, Farmland, and Park datasets.

4.1. Comparison Between Features of All Spectral Segments and Single Spectral Segments

The feature extraction method employed in this study was compared with methods outlined in the literature. SuperPoint features were extracted from the selected or constructed feature spectral segments, and these features were paired with LightGlue features. The comparison of the number of features and pairs corresponding to different feature extraction strategies for the "Lakeside Dataset" is presented in Table 1 below.

Table 1. Comparison between the number of features and feature pairs corresponding to different feature extraction strategies for the lakeside dataset.

Methods	Number of Features in the Left Data Cube	Number of Features in the Right Data Cube	Number of Successful Feature Pairs Between the Two Data Cubes
Select intermediate band [18]	1537	1663	353
Select peak signal-to-noise ratio [19]	2927	2806	622
Image fusion into one image [31]	3981	3509	897
Image fusion into one image [32]	3016	2866	742
Our method	131,117	133,511	37,876

As depicted in Table 1, selecting an intermediate band [18] for feature extraction represented a simple approach. However, the chosen intermediate spectral band lacked representativeness. Selecting the spectral segment with the highest PSNR [19] for feature extraction was representative; however, it failed to represent the features of all spectral segments in the cube. Methods involving the fusion of images into one image and subsequent feature extraction [31,32], or fusion into one image and subsequent feature extraction essentially add grayscale pixel values, resulting in a limited number of features and paired

features. The principal advantage of our method lay in preserving all features in the original spectral segments, thereby enhancing the feature pairing process. The number of features employed in the proposed method exceeded those utilized in other methods by an order of magnitude. Features from each spectral segment image were extracted, and the feature pairing results are illustrated in Figure 8.
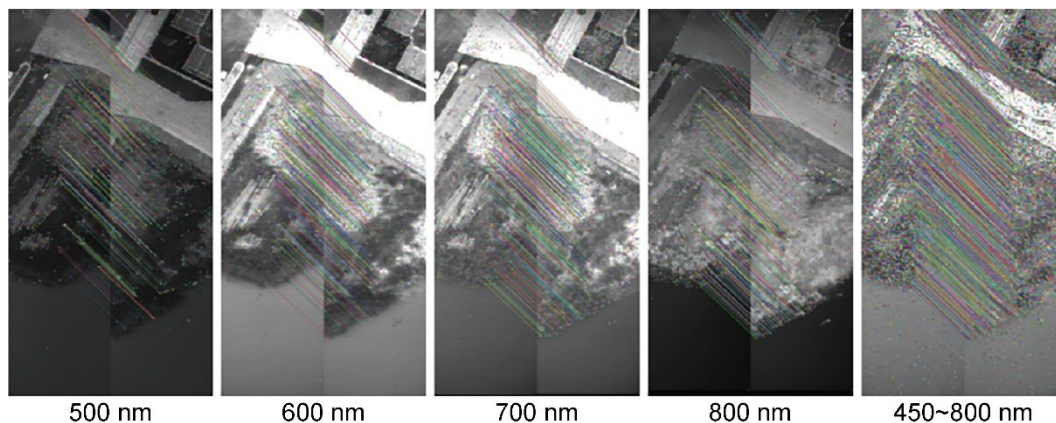

500 nm   600 nm   700 nm   800 nm   450~800 nm

Figure 8. Pairing of single and full spectral features for the Lakeside dataset.

The number and distribution of features varied across different spectral segments of the data cube. For instance, as demonstrated in the 600 and 700 nm images in Figure 8, a few features were present on the road in the overlapping area (the area with saturated brightness). Conversely, distinct paired features were detected on the overlapping roads in the 500 and 800 nm images. The image representing the 450–800 nm range illustrated the outcomes obtained by the proposed method, showcasing the pairing of all spectral segments in the data cube after feature fusion, indicative of a considerable increase in the number of features.

The results depicted in Table 2, along with the counts of left and right data cube features, and the successful matching of these features (as shown in Table 1), yielded consistent comparative outcomes. Specifically, our method demonstrated a tenfold increase in both the number of extracted features and successful matches. This verified the extraction of features from each spectral segment image, as evidenced by the feature pairing outcomes illustrated in Figure 9.

Table 2. Comparison of the number of features and feature pairs corresponding to different feature extraction strategies for the Farmland dataset.

Methods	Number of Features in the Left Data Cube	Number of Features in the Right Data Cube	Number of Successful Feature Pairs Between the Two Data Cubes
Select intermediate band [18]	1694	2290	109
Select peak signal-to-noise ratio [19]	2413	3143	121
Image fusion into one image [31]	2348	2746	96
Image fusion into one image [32]	2077	2672	88
Our method	107,049	127,143	4936

The distribution and quantity of features exhibited variations across different spectral segments of the data cube. For instance, in the Farmland dataset depicted in Figure 9, where the overlap was relatively small, the overall number of features was correspondingly diminished. Consequently, there were fewer successfully paired features within the single spectral ranges of 500 nm, 600 nm, 700 nm, and 800 nm. Our proposed method, focusing

on the 450–800 nm range, illustrated the fusion of paired feature sets across all spectral ranges, indicating a substantial increase in the number of feature pairs.
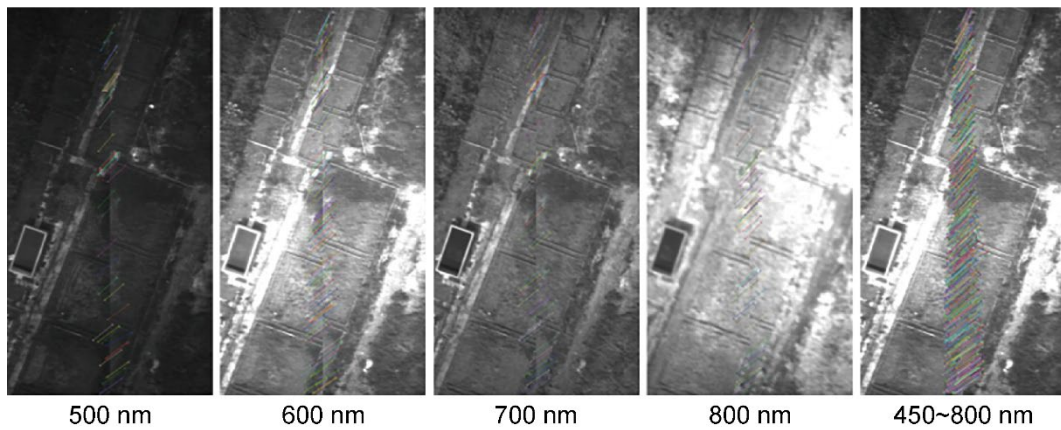

Figure 9. Pairing of single and full spectral features for the Farmland dataset.

Similarly, the findings presented in Table 3, along with the counts of left and right data cube features, and successful matching (as detailed in Table 1), confirmed a tenfold increase in both the number of extracted features and successful matches using our method. This reaffirmed the extraction of features from each spectral segment image, as demonstrated in Figure 10.

Table 3. Comparison between the number of features and feature pairs corresponding to different feature extraction strategies for the park dataset.

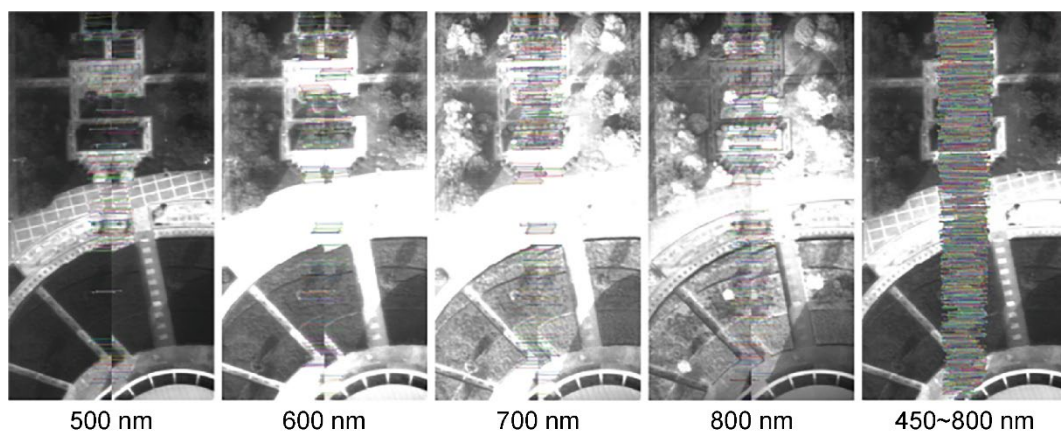| Methods	Number of Features in the Left Data Cube	Number of Features in the Right Data Cube	Number of Successful Feature Pairs Between the Two Data Cubes
Select intermediate band [18]	3099	3169	497
Select peak signal-to-noise ratio [19]	3215	3008	559
Image fusion into one image [31]	2913	3187	605
Image fusion into one image [32]	3409	3260	618
Our method	160,956	152,741	29,631

Figure 10. Pairing of single and full spectral features for the Park dataset.

The distribution and quantity of features also varied across different spectral segments of the data cube. For example, in Figure 10, images at 600 nm, 700 nm, and 800 nm exhibited saturated middle regions with no features in the overlapping areas. Conversely, multiple pairs of features were detected in the overlapping region of the 500 nm image.

Ultimately, the pairing of the spectral bands ranging from 450 nm to 800 nm in the data cube underscored a tenfold increase in the number of features.

4.2. Evaluation of Spatial Information

In information theory, entropy serves as a metric for describing information richness. Hence, entropy becomes a useful tool for evaluating the informational content within a data cube. In line with this, we formulated an entropy-based evaluation function

$$F = -\sum_{g=0}^{G} P_k(g) log_b P_k(g) \tag{8}$$

where b is typically equal to two, g represents the grayscale value of the image, G represents the maximum grayscale value of the image, k represents the defocus image sequence, and $P_k(g)$ represents the probability of the grayscale value g appearing in the kth image.

The entropy of the data cube is defined as

$$E = \sum_{n=1}^{N} F_n \tag{9}$$

where N is the total number of spectral segments in the hyperspectral data cube.

To assess the differences in informational content between the splicing strategy employed in this study and other methods, we utilized the entropy of the data cube for quantitative evaluations.

Table 4 illustrates that both the original fusion method and the method utilizing the optimal stitching technique [31,33,34] produced similar entropies for the data cubes post-stitching. However, the entropies of both methods were smaller than those of the stitching seam-finding strategy based on the hyperspectral image classification method proposed in this study.

Table 4. Entropy of data cubes after splicing using different splicing strategies.

| Datasets	Original Fusion Method	Optimal Stitching Technique [11,12]	Optimal Stitching Technique [31]	Our Method
Lakeside dataset	471.4471	472.0932	472.1074	472.8813
Farmland dataset	464.5104	464.6879	464.8157	465.0429
Park dataset	462.8703	463.7294	463.1543	464.8864

To visually discern the disparities in spatial information obtained by the splicing strategy developed in this study compared with other methods, data cubes spliced using the four methods were constructed. These methods included the original non-seam method, optimal suture technique from [33,34], optimal suture technique from [31], and our method.

From the observation in Figure 11, it was evident that the original fusion method (Figure 11a) yielded a considerable number of stitching seams in the concatenated data cube, Figure 11b was similar to the splicing seam in Figure 11c, and our method (Figure 11d) facilitated smooth transitions of the seams. The sub-images in the red boxes in Figure 11 demonstrated the stitching effect on the water surface, exhibiting increased stability even under excessive display conditions.
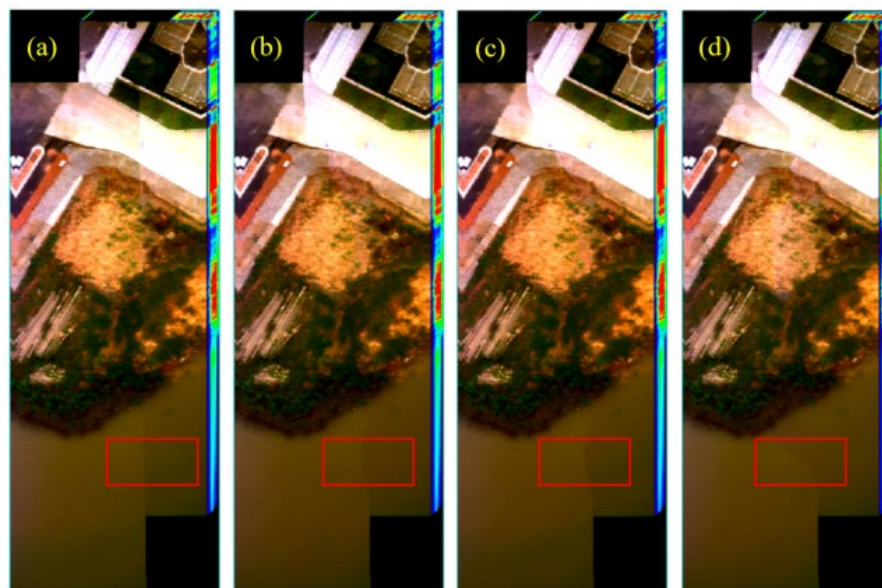
Figure 11. The stitching effect for the Lakeside dataset: (**a**) original non-seam method, (**b**) optimal suture technique from [33,34], (**c**) optimal suture technique from [31], and (**d**) our method.

From the observations in Figure 12, the following points were noted: the original fusion method (Figure 12a) resulted in noticeable stitching seams in the concatenated data cubes; slight blue ghosting was observed in the stitching of the overlapping area in Figure 12b; comparatively in Figure 12c, the blue ghosting on the stitching of the overlapping area was slightly reduced compared with (Figure 12b); and our method (Figure 12d) demonstrated minimal ghosting on the stitching of overlapping areas. The sub-images in the red boxes of Figure 12 illustrated the presence of ghosting on the stitching seam of the overlapping area.
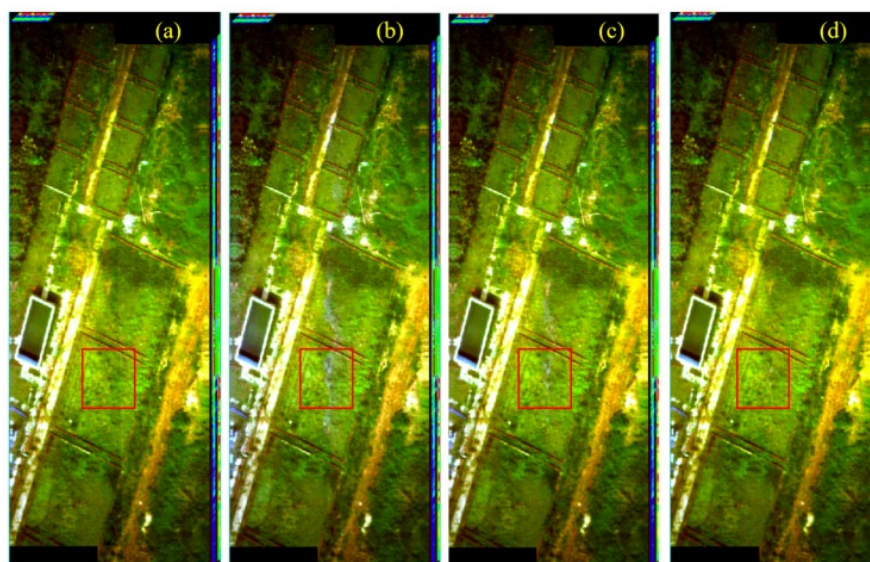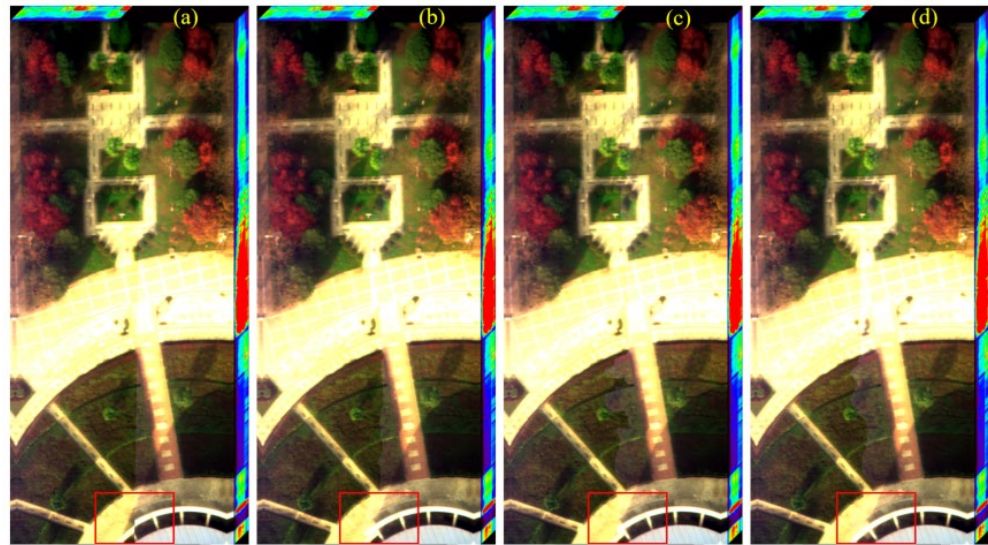

Figure 12. The stitching effect for the Farmland dataset: (**a**) original non-seam method, (**b**) optimal suture technique from [33,34], (**c**) optimal suture technique from [31], and (**d**) our method.

The following conclusions could be drawn based on Figure 13: the original fusion method (Figure 13a) led to considerable stitching seams in the spliced data cube with noticeable misalignment in the stitching of the curved building at the bottom; Figure 13b was similar to Figure 13c in the splicing joint. Additionally, there was a misalignment of splicing of the curved building at the bottom part. Our method (Figure 13d) ensured smooth and

seamless joints, with the curved buildings well-connected without any misalignment. The sub-images in the red boxes in Figure 13 indicated the presence or absence of misalignment.

Figure 13. The stitching effect for the Park dataset: (**a**) original non-seam method, (**b**) optimal suture technique from [33,34], (**c**) optimal suture technique from [31], and (**d**) our method.

4.3. Evaluations Based on Spectral Information

To assess the proposed method from the perspective of spectral information, four typical features were chosen in the overlapping area, and their spectral curves were extracted using the same control points as those in the left data cube. Spectral angle mapping (SAM) was employed to determine the similarity between the two arrays [31]. SAM computes similarity values based on the overall resemblance of spectral curves, treating each pixel's spectrum in an image as a high-dimensional vector. The similarity between spectra was gauged by calculating the angle between the two vectors. A smaller angle indicated a greater similarity between the two spectra. Denoting the target and reference spectra as x and y, respectively, each with a length of *n*, we could define

$$SAM(x,y) = \arccos \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2}} \tag{10}$$

where $SAM(x,y)$ is the spectral angle between the two spectra. The modes were two spectral vectors.

As depicted in Figure 14, six points were selected in the overlapping area of the left and right data cubes across three datasets, corresponding to (a) the road surface, (b) stone, (c) grassland, (d) water body, (e) crops, and (f) trees.

Figure 15 displays the spectral curves of four points extracted for data cubes concatenation using different methods, with the reference spectrum being the spectral curve extracted from the corresponding ground object in the left data cube. SAM was then calculated for spectral curves extracted by different methods and reference spectral curves. From the graph, it was evident that the method proposed in this article exhibited a higher degree of overlap between the extracted spectral curve and the reference curve. Notably, for point (c), the spectral curve of the grassland demonstrated the highest similarity from 700 nm to 800 nm. Additionally, from an overall perspective, the differences in spectral curves for (a), (b), (d), (e), and (f), were not pronounced. However, based on numerical analyses of the spectral curves, our proposed method yielded the SAM values of 0.0125, 0.0168, 0.0210, 0.0279, 0.0286, and 0.0283, respectively, which were smaller than those calculated by other

methods. Therefore, our proposed stitching seam-finding method based on hyperspectral image classification also offered certain advantages in spectral performance following data cube stitching.
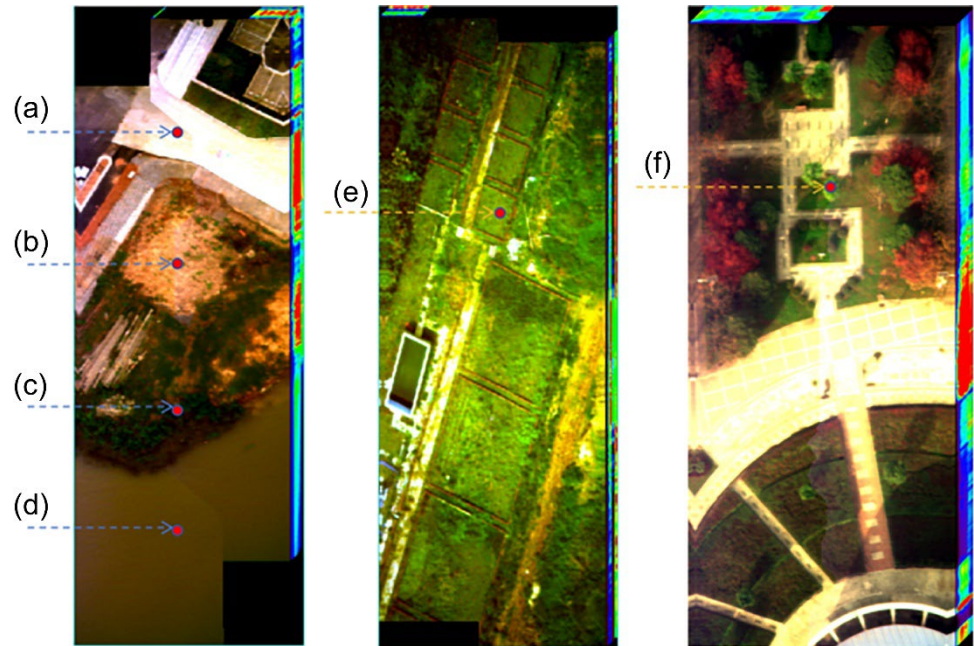

Figure 14. Typical land features in hyperspectral data cube images: (**a**) a road, (**b**) stone, (**c**) grass, (**d**) water, (**e**) crops, and (**f**) a tree.
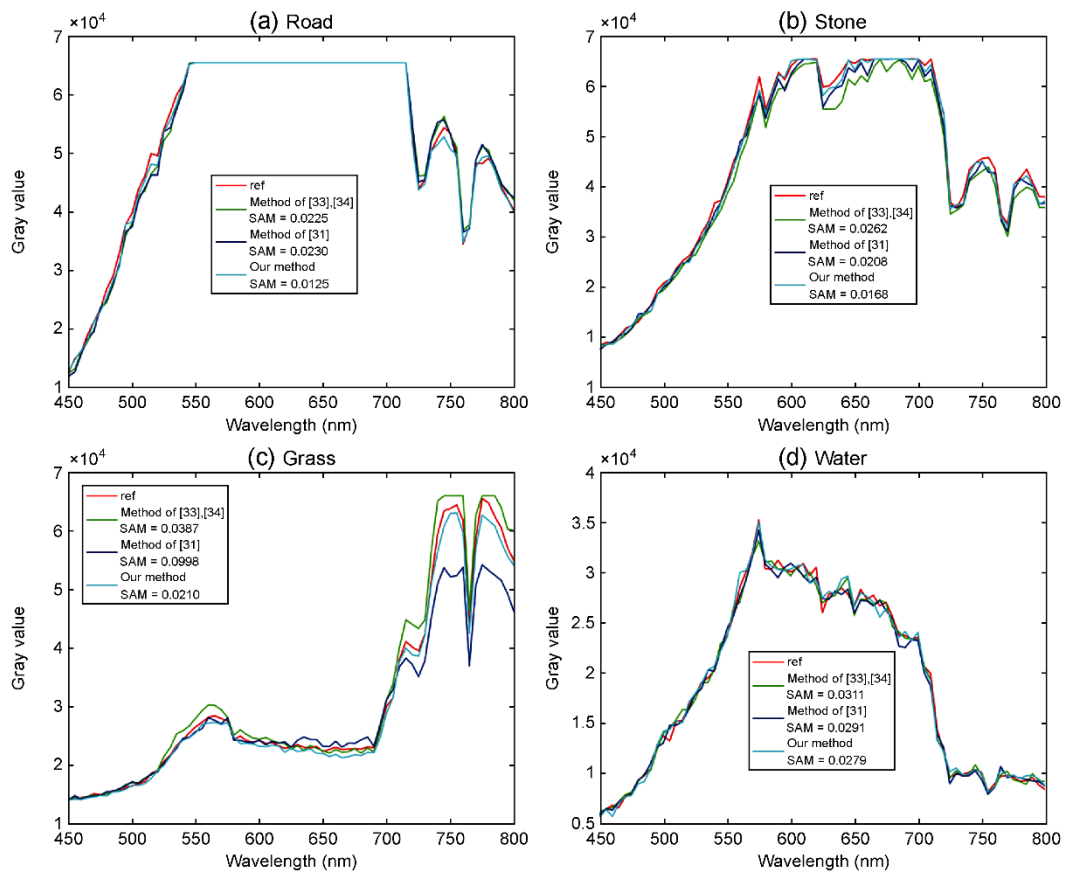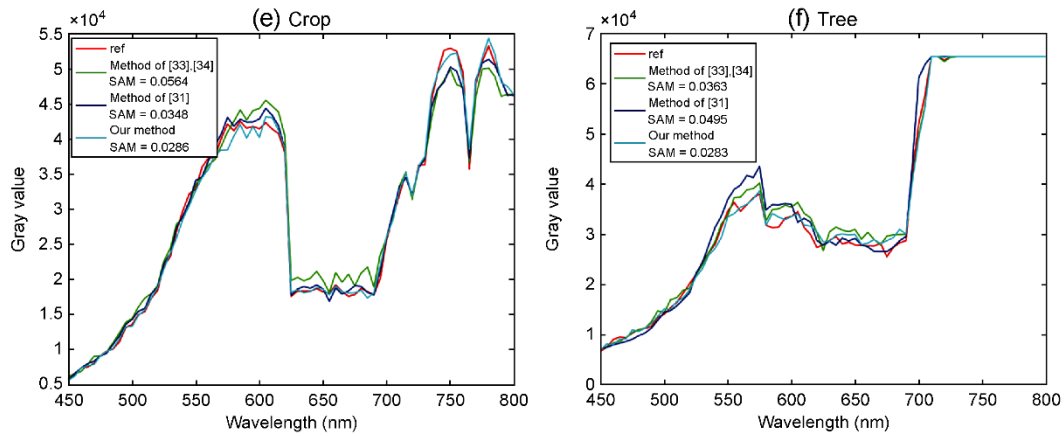

Figure 15. *Cont.*

Figure 15. Spectral angle mapping (SAM) results for the spectral curves of the six points shown in Figure 14.

4.4. Computational Complexity Analysis

In terms of algorithmic complexity, compared with the method proposed in this paper, the calculation process of the method [18] was to select the image of the 88th band in the middle of 176 bands from the two data cubes as the feature extraction image, including feature matching based on the SIFT and mTopKRP algorithms, as well as robust elastic warp and multiband fusion to obtain a hyperspectral panorama. The implementation process of the method [19] was as follows: One band was selected as the reference using the PSNR estimation, Exact SuperPoint feature points, and Match feature points, and false matches were removed using LAF, image alignment using robust elastic warp, and adaptive bundle adjustment. All the bands were fused to obtain the final panoramic hyperspectral image.

The implementation process of method [31] was as follows. First, band fusion was performed; the SuperPoint network was used to extract feature points, and the SuperGlue network was used to obtain matching point pairs. The transformation matrix was estimated by robust elastic image warping. The input HSIs were warped into the coordinate system of the stitching result. The process implemented in methods [33] and [34] involved inputting two data cubes composed of 176 spectral segments, implementing image registration and alignment using the mTopKRP+REW method, optical coastline detection using the Graph cut method, and the stitching HSI. The implementation process of the method proposed in this article is shown in Figure 4. To fully utilize the spectral information of the data cube, the implementation process was more complex than the comparative methods.

Specifically, to assess the computational complexity of the methods proposed in this article, we conducted a statistical analysis of the runtime of all methods using the Lakeside dataset. Table 5 presents a comparison of the time taken to complete the stitching of the Lakeside dataset using different stitching strategies.

Table 5. Comparison of the time taken to complete the stitching of the Lakeside dataset using different stitching strategies (in s).

/	Select Intermediate Band [18]	Select Peak Signal-to-Noise Ratio [19]	Image Fusion Into One Image [31]	Image Fusion Into One Image [33]	Our Method
Stitching time	127.58	228.25	162.20	289.35	952.57 (118.23)

According to the results in Table 5, it was evident that our proposed method required the longest time to complete the stitching process overall. However, in the experimental process, the time required for each stage of our method was 846.26 + 19.07 + 87.24 = 952.57 s. Among them, the time required for data cube feature extraction, all feature fusion,

and all feature matching was 846.26 s. The K-means classification of hyperspectral data cubes consumed 19.07 s, while 87.24 s were spent on finding and applying the stitching seam to all spectral segments. If only one spectral image was selected for using our proposed method, the entire stitching process would take approximately (846.26/71 bonds) + 19.07 + 87.24 $\approx$ 118.23 s. Despite the extended duration required by our proposed strategy to complete the dataset stitching, it was noteworthy that even in the scenarios of low-imaging quality, our method could achieve efficient and accurate cube stitching of the hyperspectral data across various scenes.
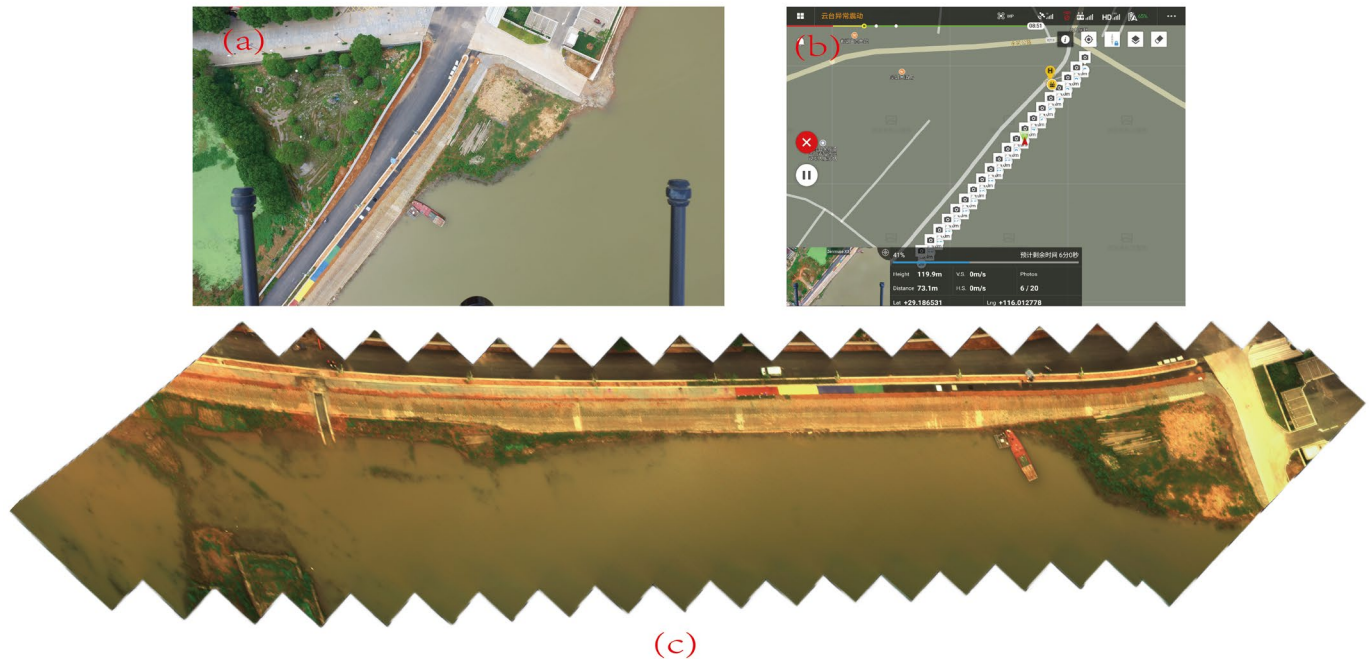
4.5. Large-Field Stitching Results

To validate the method proposed in this article, an outdoor drone flight test (110°44′58.6″ E, 30°57′21.4″ N) was added. Hyperspectral data from five waypoints captured using the Park dataset were used and the principle diagram of the hyperspectral data cube stitching proposed in this article was used to complete the stitching of the hyperspectral remote sensing data from the five data cubes. Additionally, extraction was performed (red: 650 nm, green: 530 nm, blue: 480 nm) to synthesize a pseudo-color image, as shown in Figure 16.
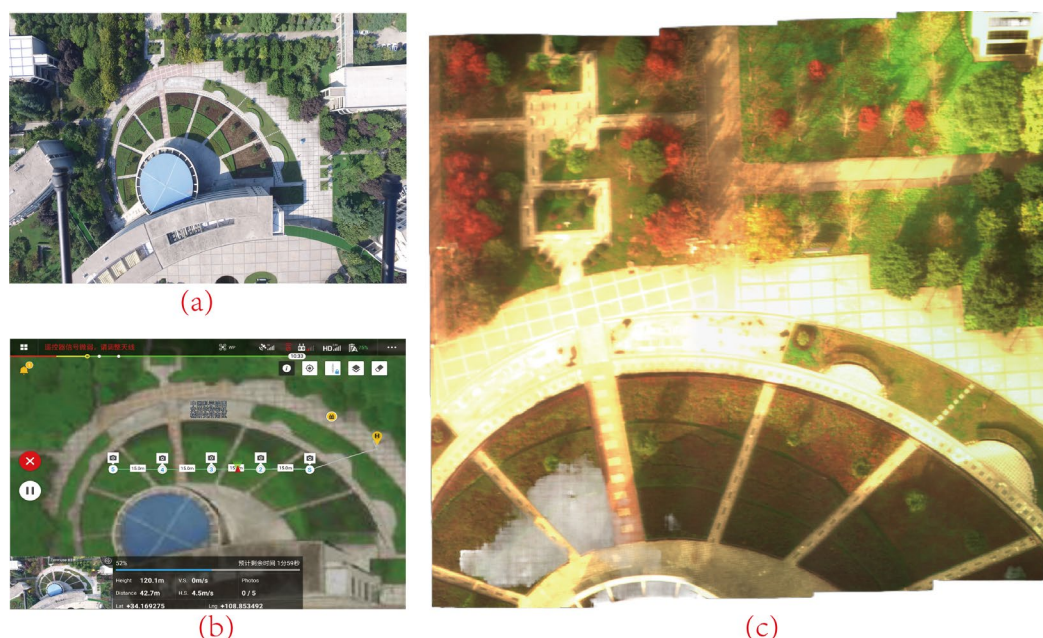
Figure 16. Large field stitching of the UAV hyperspectral remote sensing images taken from the waypoints. (**a**) Ground images captured by red–green–blue (RGB) cameras on the UAV. (**b**) The setting interface for hyperspectral imaging in UAV waypoint planning. (**c**) Splicing results of the hyperspectral remote sensing data from five waypoints.

To validate the method proposed in this article, a second outdoor drone flight test (116°0′46.3″ E, 29°11′12.5″ N) was added. Hyperspectral data from 19 waypoints, captured using the Lakeside dataset, were used and the principle diagram of the hyperspectral data cube stitching proposed in this article was used to complete the stitching of the 19 hyperspectral remote sensing data cubes. Additionally, extraction was performed (red: 650 nm, green: 530 nm, blue: 480 nm) to synthesize a pseudo-color image, as shown in Figure 17 below.

Figure 17. Large-field stitching of the UAV hyperspectral remote sensing images. (**a**) Ground images captured by RGB cameras mounted on the UAV. (**b**) Interface for the hyperspectral imaging setup in UAV waypoint planning. (**c**) Results of the stitching of 19 hyperspectral data cubes.

5. Discussion

Based on the experimental results presented above, several facts were observed.

Based on the Lakeside, Farmland, and Park datasets, combined with Tables 1–3 and the results shown in Figures 7–9, in the feature extraction stage of hyperspectral image stitching in references [18,19,31,32], the selected feature extraction objects were ultimately one image; one image cannot represent the entire data cube. The number of extracted features from one image was small, making it difficult to represent all the feature information of a data cube. The superiority of the proposed method can also be deduced from the number of features. That is, this study proposed a new multi-feature fusion strategy that used the features of all the spectral segments as the overall features of the data cube. These fused features could comprehensively reflect the characteristics of each spectral segment of the data cube.

The listings in Table 2 showed that the mean entropies of the concatenated data cubes for the three datasets using the original fusion method and the optimal stitching techniques [31,33,34] were similar, but were all smaller than the values yielded by the stitching seam-finding strategy based on the hyperspectral image classification proposed herein. This is because—in addition to the original fusion method, and among the two types of seam-finding methods mentioned above—only the first method [33,34] selected the seam line with the maximum SSIM for a specific spectral segment, and could not achieve the maximum SSIM for the spliced images of the other spectral segments. In the second method [31], the essence of seam selection was to also find the optimal seam, which was based on the constructed image and did not fully consider the situation of all the spectral segments. Furthermore, this study proposed a stitching seam-search strategy based on hyperspectral image classification. This method also overcame the drawback of using only a single spectral segment to search for seam joints without fully utilizing the spectral information, resulting in inaccurate seam joints.

This study was associated with limitations. By analyzing the complexity of the proposed method based on the feature fusion strategy and additional hyperspectral im-

age classification, the stitching time was increased considerably compared with those of other methods.

Future research will continue to optimize the algorithm and reduce its time complexity. In addition, from a hardware perspective, hardware implementation algorithms can be used to implement the proposed splicing method on FPGA [46,47].

In addition, the future research plan is to achieve large field-of-view stitching of the hyperspectral remote sensing images obtained by a UAV-mounted spectral scanning spectrometer. As part of this plan, we aspire to identify ways to improve seam transition after stitching multiple data cubes.

6. Conclusions

To achieve large-field stitching of the hyperspectral remote sensing images collected using a UAV-mounted spectral scanning spectrometer, this study introduced a stitching method based on a multi-feature fusion strategy and a seam-finding approach that utilized hyperspectral image classification. The proposed method addressed the underutilization of spectral dimension information prevalent in current data cube stitching techniques. At the feature extraction stage, a multi-feature fusion strategy was devised to overcome the limitation of using single spectral images to compensate for the insufficient number of features in the data cube. This strategy involved the utilization of all the spectral features as comprehensive features of the data cubes. For the seam selection stage, a seam-finding strategy based on spectral image classification was employed to uphold the stability of the spectral information while preserving identical features in the overlapping area. The introduction of boundary lines between different features into the stitching line facilitated the experimental analysis of both spatial and spectral information contained in the concatenated data cube. The performance evaluation of the proposed strategy demonstrated its superiority over existing methods for stitching hyperspectral data cube images.

Data Availability Statement: All data generated or analyzed during this study are included in this published article.

References

1. Jakob, S.; Zimmermann, R.; Gloaguen, R. The need for accurate geometric and radiometric corrections of drone-borne hyperspectral data for mineral exploration: MEPHySToA toolbox for pre-processing drone-borne hyperspectral data. *Remote Sens.* **2017**, *9*, 88. [CrossRef]
2. Honkavaara, E.; Eskelinen, M.A.; Polonen, I.; Saari, H.; Ojanen, H.; Mannila, R.; Holmlund, C.; Hakala, T.; Litkey, P.; Rosnell, T.; et al. Remote sensing of 3-D geometry and surface moisture of a peat production area using hyperspectral frame cameras in visible to short-wave infrared spectral ranges onboard a small unmanned airborne vehicle (UAV). *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 5440–5454. [CrossRef]

3. Nevalainen, O.; Honkavaara, E.; Tuominen, S.; Viljanen, N.; Hakala, T.; Yu, X.; Hyyppa, J.; Saari, H.; Polonen, I.; Imai, N.N.; et al. Individual tree detection and classification with UAV-based photogrammetric point clouds and hyperspectral imaging. *Remote Sens.* **2017**, *9*, 185. [CrossRef]

4. Zhang, T.; Bi, Y.; Xuan, C. Convolutional transformer attention network with few-shot learning for grassland degradation monitoring using UAV hyperspectral imagery. *Int. J. Remote Sens.* **2024**, *45*, 2109–2135. [CrossRef]

5. Lakshmi, G.J.; Bodapati, V.; Baji, S.S.; Panuganti, T.S. Water Quality Monitoring Using Remote-sensing. In Proceedings of the 2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), Greater Noida, India, 28–30 April 2023; pp. 624–628.

6. Zhao, E.; Qu, N.; Wang, Y.; Gao, C. Spectral reconstruction from thermal infrared multispectral image using convolutional neural network and transformer joint network. *Remote Sens.* **2024**, *16*, 1284. [CrossRef]

7. Mo, Y.; Wei, X.; Kang, X.; Zhang, S.; Li, S. Seam-cutting Based Unmanned Aerial Vehicle Hyperspectral Image Stitching. In Proceedings of the IEEE International Geoscience Remote Sensing Symposium IGARSS 2021, Brussels, Belgium, 11–16 July 2021; pp. 8253–8256.

8. Fang, J.; Wang, X.; Zhu, T.; Liu, X.; Zhang, X.; Zhao, D. A novel mosaic method for UAV-BASED hyperspectral images. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Yokohama, Japan, 28 July–2 August 2019; Volume 2019, pp. 9220–9223. [CrossRef]

9. Chen, P.; Xu, X. A comparison of photogrammetric software packages for mosaicking unmanned aerial vehicle (UAV) images in agricultural application. *Acta Agron. Sin.* **2020**, *46*, 1112–1119. [CrossRef]

10. Xie, W.; Liu, T.; Gu, Y. Intrinsic hyperspectral image recovery for UAV strips stitching. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5527013. [CrossRef]

11. Yi, L.-N.; Xu, X.; Zhang, G.-F.; Ming, X.; Sha, L.-Y. Light and small UAV hyperspectral image mosaicking. *Spectrosc. Spectral Anal.* **2019**, *39*, 1885–1891.

12. Xiong, J.; Li, F.; Long, F.; Xu, Y.; Wang, S.; Xu, J.; Ling, Q. Spatially-varying Warping for Panoramic Image Stitching. In Proceedings of the 34th Chinese Control and Decision Conference (CCDC), Hefei, China, 15–17 August 2022; pp. 575–580.

13. Li, L.; Yao, J.; Xie, R.; Li, J. Edge-enhanced optimal seamline detection for orthoimage mosaicking. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 764–768. [CrossRef]

14. Yi, L.; Chen, J.M.; Zhang, G.; Xu, X.; Ming, X.; Guo, W. Seamless mosaicking of UAV-based push-broom hyperspectral images for environment monitoring. *Remote Sens.* **2021**, *13*, 4720. [CrossRef]

15. Chung, K.-L.; Chang, W.-T. Centralized RANSAC-based point cloud registration with fast convergence and high accuracy. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 5431–5442. [CrossRef]

16. Ma, J.; Zhao, J.; Tian, J.; Yuille, A.L.; Tu, Z. Robust point matching via vector field consensus. *IEEE Trans. Image Process.* **2014**, *23*, 1706–1721. [CrossRef]

17. Ma, J.; Jiang, J.; Zhou, H.; Zhao, J.; Guo, X. Guided locality preserving feature matching for remote sensing image registration. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4435–4447. [CrossRef]

18. Zhang, Y.; Wan, Z.; Jiang, X.; Mei, X. Automatic stitching for hyperspectral images using robust feature matching and elastic warp. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 3145–3154. [CrossRef]

19. Zhang, Y.; Mei, X.; Ma, Y.; Jiang, X.; Peng, Z.; Huang, J. Hyperspectral panoramic image stitching using robust matching and adaptive bundle adjustment. *Remote Sens.* **2022**, *14*, 4038. [CrossRef]

20. Xiao, G.; Luo, H.; Zeng, K.; Wei, L.; Ma, J. Robust feature matching for remote sensing image registration via guided hyperplane fitting. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5600714. [CrossRef]

21. Jiang, Z.; Zhang, Z.; Liu, J.; Fan, X.; Liu, R. Multi-spectral Image Stitching via Spatial Graph Reasoning. In Proceedings of the MM '23: 31st ACM International Conference on Multimedia, Ottawa, ON, Canada, 29 October–3 November 2023; pp. 472–480.

22. Zaragoza, J.; Chin, T.-J.; Brown, M.S.; Suter, D. As-projective-as-possible image stitching with moving DLT. In Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Portland, OR, USA, 23–28 June 2013; pp. 2339–2346.

23. Li, J.; Wang, Z.; Lai, S.; Zhai, Y.; Zhang, M. Parallax-tolerant image stitching based on robust elastic warping. *IEEE Trans. Multimed.* **2018**, *20*, 1672–1687. [CrossRef]

24. Li, N.; Xu, Y.; Wang, C. Quasi-homography warps in image stitching. *arXiv* **2017**, arXiv:1701.08006. [CrossRef]

25. Yang, L.; Tian, B.; Dang, J. Research on image stitching algorithm based on deep learning: A survey. *Appl. Res. Comput.* **2024**, *41*, 1930–1939.

26. Cheng, S.; Yang, F.; Chen, Z.; Yuan, N.; Tao, W. Deep seam prediction for image stitching based on selection consistency loss. *arXiv* **2023**, arXiv:2302.05027.

27. Jia, Q.; Li, Z.; Fan, X.; Zhao, H.; Teng, S.; Ye, X.; Latecki, L.J. Leveraging Line-point Consistence to Preserve Structures for Wide Parallax Image Stitching. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Electrical Network, Nashville, TN, USA, 19–25 June 2021; pp. 12181–12190.

28. Liao, T.; Chen, J.; Xu, Y. Quality evaluation-based iterative seam estimation for image stitching. *Signal Image Video Process.* **2019**, *13*, 1199–1206. [CrossRef]

29. Li, J.; Zhou, Y. Automatic Color Image Stitching Using Quaternion Rank-1 Alignment. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 19–24 June 2022; pp. 19688–19697.

30. Nie, L.; Lin, C.; Liao, K.; Liu, S.; Zhao, Y. Unsupervised deep image stitching: Reconstructing stitched features to images. *IEEE Trans. Image Process.* **2021**, *30*, 6184–6197. [CrossRef]

31. Mo, Y.; Kang, X.; Duan, P.; Li, S. A robust UAV hyperspectral image stitching method based on deep feature matching. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5517514. [CrossRef]

32. Li, S.; Yin, Q.; Hu, Y.; Gong, C.-L. A push-sweep hyperspectral aerial image mosaic method based on SPHP. *J. Infrared Millim. Waves* **2021**, *40*, 64–73.

33. Peng, Z.; Ma, Y.; Zhang, Y.; Li, H.; Fan, F.; Mei, X. Seamless UAV hyperspectral image stitching using optimal seamline detection via graph cuts. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5512213. [CrossRef]

34. Peng, Z.; Ma, Y.; Mei, X.; Huang, J.; Fan, F. Hyperspectral image stitching via optimal seamline detection. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 5507805. [CrossRef]

35. Hejazifar, H.; Khotanlou, H. Fast and robust seam estimation to seamless image stitching. *Signal Image Video Process.* **2018**, *12*, 885–893. [CrossRef]

36. Jiang, Q.; Qiu, Y.; Wen, Y.; Wang, H.; Xu, W. Design of data acquisition system for AOTF polarization spectral imaging instrument. *Infrared Las. Eng.* **2012**, *41*, 218–222.

37. Liu, H.; Yu, T.; Hu, B.; Hou, X.; Zhang, Z.; Liu, X.; Liu, J.; Wang, X.; Zhong, J.; Tan, Z.; et al. UAV-borne hyperspectral imaging remote sensing system based on acousto-optic tunable filter for water quality monitoring. *Remote Sens.* **2021**, *13*, 4069. [CrossRef]

38. Wang, J.; Zhao, D.; Du, X.; Wang, Q.; Lin, Z.; Cheng, L.; Li, Z. Design and experiment of hyper-spectral polarization imaging system based on AOTF. *Infrared Las. Eng.* **2017**, *46*, 138002. [CrossRef]

39. Sun, G.; Huang, W.; Chen, P.; Gao, S.; Wang, X. Advances in UAV-based multispectral remote sensing applications. *Trans. Chin. Soc. Agric. Mach.* **2018**, *49*, 1–17. [CrossRef]

40. Song, Y.; Zhang, D.; Tang, L.; Li, C.; Ma, L. An approach to stitch UAV image with low overlap based on ASIFT algorithm. *Remote Sens. Technol. Appl.* **2015**, *30*, 725–730.

41. Ren, X.; Sun, M.; Zhang, X.; Liu, L. A simplified method for UAV multispectral images mosaicking. *Remote Sens.* **2017**, *9*, 962. [CrossRef]

42. DeTone, D.; Malisiewicz, T.; Rabinovich, A. SuperPoint: Self-supervised Interest Point Detection and Description. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–22 June 2018; pp. 337–349.

43. Lindenberger, P.; Sarlin, P.E.; Pollefeys, M. LightGlue: Local Feature Matching at Light Speed. In Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision, Vancouver, BC, Canada, 18–22 June 2023; pp. 17581–17592.

44. Sarlin, P.-E.; DeTone, D.; Malisiewicz, T.; Rabinovich, A. SuperGlue: Learning Feature Matching With Graph Neural Networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Electrical Network, Seattle, WA, USA, 13–19 June 2020; pp. 4937–4946.

45. Liu, H.; Hou, X.; Hu, B.; Yu, T.; Zhang, Z.; Liu, X.; Liu, J.; Wang, X.; Zhong, J.; Tan, Z. Image blurring and spectral drift in imaging spectrometer system with an acousto-optic tunable filter and its application in UAV remote sensing. *Int. J. Remote Sens.* **2022**, *43*, 6957–6978. [CrossRef]

46. Jia, Y.; Wang, R.; Jiang, X. A real-time image stitching and fusion algorithm circuit design based on FPGA. *Electronics* **2024**, *13*, 271. [CrossRef]

47. Palacios, P.; Bascones, D.; Gonzalez, C.; Mozos, D. A real-time FPGA Implementation of the LCMV algorithm for target classification in hyperspectral images using LDL decomposition. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5524814. [CrossRef]