Communication

Fast Automatic Registration of UAV Images via Bidirectional Matching

Xin Luo [1], Zuqi Wei [1], Yuwei Jin [2], Xiao Wang [1], Peng Lin [1], Xufeng Wei [1] and Wenjian Zhou [3,*]

1. Yangtze Delta Region Institute (HuZhou), University of Electronic Science and Technology of China, Huzhou 313099, China; luoxin@uestc.edu.cn (X.L.); wangxiao29@std.uestc.edu.cn (X.W.); 2019070403014@std.uestc.edu.cn (P.L.); xufengw@std.uestc.edu.cn (X.W.)
2. School of Electrical and Information Engineering, Panzhihua University, Panzhihua 617000, China; jinyuwei@alu.uestc.edu.cn
3. School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
* Correspondence: zhouwenjian@uestc.edu.cn; Tel.: +86-1801-062-2207

Abstract: Image registration plays a vital role in the mosaic process of multiple UAV (Unmanned Aerial Vehicle) images acquired from different spatial positions of the same scene. Aimed at the problem that many fast registration methods cannot provide both high speed and accuracy simultaneously for UAV visible light images, this work proposes a novel registration framework based on a popular baseline registration algorithm, ORB—the Oriented FAST (Features from Accelerated Segment Test) and Rotated BRIEF (Binary Robust Independent Elemental Features) algorithm. First, the ORB algorithm is utilized to extract image feature points fast. On this basis, two bidirectional matching strategies are presented to match obtained feature points. Then, the PROSRC (Progressive Sample Consensus) algorithm is applied to remove false matches. Finally, the experiments are carried out on UAV image pairs about different scenes including urban, road, building, farmland, and forest. Compared with the original version and other state-of-the-art registration methods, the bi-matching ORB algorithm exhibits higher accuracy and faster speed without any training or prior knowledge. Meanwhile, its complexity is quite low for on-board realization.

Keywords: UAV; image registration; ORB; point feature; bidirectional matching

1. Introduction

Because of its flexibility, high efficiency, and low cost, UAV (Unmanned Aerial Vehicle) remote sensing technology has gradually emerged in many fields, such as accurate agriculture, resources investigation, environment management, and disaster monitoring [1–3]. How to yield high-precision registered UAV images quickly has become an inevitable challenge to the wide application of UAV technology [4–9]. However, the high resolutions of UAV images have a great influence on the detection and matching of image feature points. Meanwhile, there is a series of parameters in collecting UAV images, such as flight speed, flight height, and weather conditions, which make detection and matching of image feature points more difficult to achieve. Hence, the automatic registration of UAV remote sensing images is worth studying comprehensively.

Recently, the issue of image registration has received considerable intensive attention from scholars. For instance, the SIFT (Scale Invariant Feature Transform) operator can obtain a sort of feature invariant to image resolution, rotation, and scaling [10]. However, the SIFT algorithm will yield 128-dimensional features, which is heavily time-consuming. Hence, the feature dimension in the SURF (Speeded-Up Robust Features) algorithm is reduced to 64 through wavelet transform, which efficiently enhances the registration speed [11,12]. Moreover, focused on matching different features, Zhang et al. combined the image enhancement technology with the SURF algorithm to obtain better feature points

and matching efficiencies [13]. Wei et al. divided an image into super-pixels to calculate the information entropy of each region [14]. The redundancies in feature points are diminished according to the values of their information entropy. Zhang et al. proposed a threshold evaluation strategy to extract SURF rough matching points and employed the RANSAC (Random Sample Consensus) algorithm to eliminate gross errors [15]. Furthermore, in order to constrain the regions for fine matching, a similarity function of Delaunay triangles constructed by Harris points is introduced to obtain matched triangle pairs. At present, ORB (the Oriented FAST and Rotated BRIEF) is a fast local feature detection operator [16], which is improved based on the FAST (Features from Accelerated Segment Test) [17] and BRIEF (Binary Robust Independent Elemental Features) algorithm [18]. It overcomes the lack of scale and rotation invariances in FAST and is faster than most classical image registration algorithms. However, since it has the disadvantage of low accuracy, this work attempts to improve the comprehensive performances of ORB on UAV images. The main contributions of this work are presented below.

a. The ORB operator is utilized to extract image feature points since many fast registration methods cannot simultaneously provide high speed and accuracy for UAV images. Its complexity is quite low for on-board realization.

b. The KNN-based search and the Hamming distance are used in the initial matching of feature points; two bi-directional matching strategies are designed for their fine matching; and the progressive consistency algorithm (PROSAC) is exploited to remove false matches and fit a geometric transformation model to enhance registration accuracy.

c. The validation experiments were carried out on the UAV images from five typical scenes, which were acquired by different UAV image sensors. The results indicate that the proposed algorithm can achieve higher accuracy and faster speed than some existing fast registration algorithms.

The remainder of this article is organized as follows. Section 2 describes an improved ORB registration method. After images are down-sampled to speed up feature matching, an improved ORB operator and a bi-directional matching strategy are utilized to detect and match feature points in UAV images. After the initial matching, feature point pairs are obtained, and the PROSRC (Progressive Sample Consensus) algorithm is performed to screen out mismatched points to obtain an accurately matched point-pair set and final transform parameters [19]. In Section 3, our proposed algorithm is compared with some existing registration methods using UAV images acquired from different scenarios and sensors. The conclusion and discussion are presented in Section 4.

2. Methodology

Because of their high resolutions, it usually takes a long time for common image registration algorithms to process UAV images. However, there are often real-time requirements in many applications of UAV images. Therefore, based on an improved ORB operator, a bi-directional matching strategy is combined with PROSAC to pair extracted feature points, which can enhance the accuracies of the proposed registration algorithm. The registration procedure of our algorithm is illustrated in Figure 1.

2.1. Image Preprocessing
2.1.1. Down-Sampling

On account of the high resolutions of UAV images, an appropriate scaling factor is chosen to adjust the sizes of input images. If the initial size of an image is $M \times N$, the width and height of the image are reduced to $1/n$ of the original one after being down-sampled by n. That is, the size of the obtained image becomes $(M \times N)/n$, where the pixel values of each $n \times n$ region in the original image are replaced by the mean pixel value of the region. This operation can reduce the speeds of image loading and processing, which is beneficial to the real-time performances of registration algorithms.
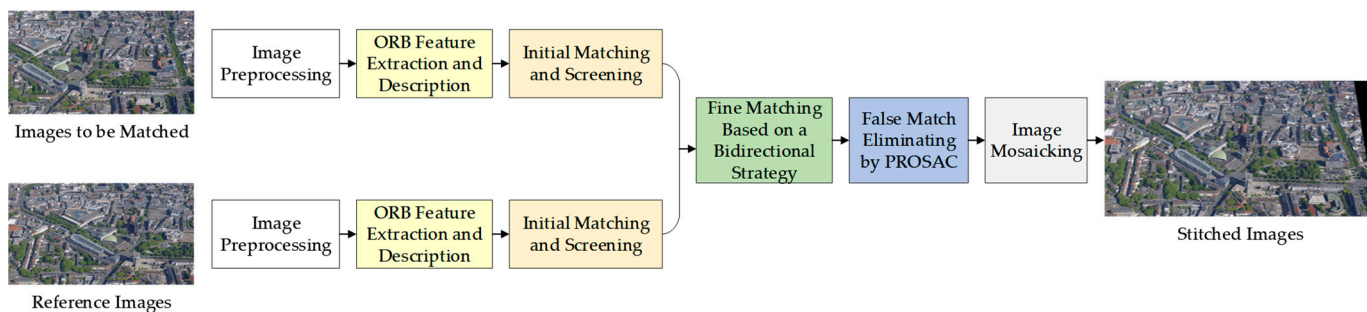
Figure 1. The flow chart of the proposed fast registration algorithm for UAV (Unmanned Aerial Vehicle) images.

2.1.2. Graying

It is known that the structure and texture characteristics of visible images usually will not vary with optical bands. For the sake of processing speeds, grayscale images are used as the basis of image registration in this work, and input RGB images are grayed according to the importance of three optical bands. The intensity of each pixel is generated from a weighted average of three different components. A graying instance of UAV images is given in Figure 2. Considering the sensitivity of human eyes to color information, the intensity of an arbitrary pixel can be denoted as

$$f(i,j) = 0.30R(i,j) + 0.59G(i,j) + 0.11B(i,j) \tag{1}$$

where $R(i,j)$, $G(i,j)$, and $B(i,j)$ represent the three components at an arbitrary piont (i,j) in an input image [20].

(**a**)   (**b**)

Figure 2. A graying instance of UAV images: (**a**) a visible image; (**b**) the grayed image.

2.2. ORB Brief

To ensure scale invariances, a Gaussian pyramid and patch centroid calculation based on FAST are combined in ORB. Moreover, an improvement on the BRIEF algorithm for rotation invariance is designed for constructing feature descriptors of ORB [21–23].

2.2.1. Feature Point Extraction

In terms of feature point detection, the kernel method of ORB is based on FAST [24–26]. If the difference in the gray level between the central pixel and one of its neighbors is significant enough, the point will be regarded as a feature point. This detection method of feature points not only preserves image features as much as possible but also greatly reduces computational complexity. It will bring about a significant improvement in the

speed of the ORB algorithm. For an arbitrary detected feature point, the moment of its neighbor patch m_{pq} is defined as

$$m_{pq} = \sum_{x,y} x^p y^q I(x,y), \quad p,q \in \{0,1\} \tag{2}$$

where $I(x, y)$ is the intensity of the pixel at (x, y). Then, the centroid coordinate C can be determined by

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \tag{3}$$

Suppose that the feature point is O. The angle of the vector $\overrightarrow{OC}$ θ is regarded as the orientation of the feature point, as given below.

$$\theta = \arctan2 \left(\frac{m_{01}}{m_{10}} \right) \tag{4}$$

where arctan2 is the quadrant-aware arctan [24].

2.2.2. Feature Descriptor Construction

After the feature points in images are found, an improved BRIEF algorithm, rBRIEF (rotation-aware BRIEF), is adopted in ORB to describe the features of feature points [27–29]. Binary coding is employed by BRIEF to construct feature description vectors, which is very helpful for accelerating calculation. Since BRIEF does not possess rotation invariance, the specific improvements on the BRIEF algorithm in ORB are as follows [24].

Give an $M \times M$ pixel smoothed patch P. A binary testing $\tau(P; u, v)$ can be defined as

$$\tau(P; u, v) = \begin{cases} 1 & P(u) < P(v) \\ 0 & otherwise \end{cases} \tag{5}$$

where $P(u)$ and $P(v)$ are the intensities of the pixels at two points u and v, respectively. For n pairs of points around a feature point, the generated feature descriptor $f_n(P)$ is an n-dimensional binary string, as follows.

$$fn(P) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(P; u_i, v_i) \tag{6}$$

According to the orientation of feature points, a corresponding matrix S is constructed as given in the following Equation.

$$S = \begin{bmatrix} u_1, u_2, \cdots, u_n \\ v_1, v_2, \cdots, v_n \end{bmatrix} \tag{7}$$

Since the orientation of the feature point is θ, the corresponding rotation matrix R_θ is defined by

$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{8}$$

Then, if the matrix S is transformed by the rotation matrix R_θ, a new feature description matrix S_θ is obtained, which can be derived as below.

$$S_\theta = R_\theta S = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} u_1, u_2, \cdots, u_n \\ v_1, v_2, \cdots, v_n \end{bmatrix} \tag{9}$$

Consequently, the improved descriptor $g_n(P, \theta)$ can be expressed as

$$g_n(P, \theta) = f_n(P) | (u_i, v_i) \in S_\theta \tag{10}$$

2.3. Initial Matching and Screening

The initial matched point pairs are yielded by comparing the extracted feature points in the image to be registered and its reference image. In practice, the comparison of image point features is realized based on the similarities of feature points. The Hamming distance is chosen to describe the similarities between different points in this research [30–32]. To enhance the efficiency of initial matching, a fast search based on the K-Nearest Neighbor (KNN) algorithm is adopted in this work [33,34]. Unlike common violent matching methods, it searches for a corresponding point with the highest similarity for a point to be registered in a determined region of the reference image. Their Hamming distances to the point to be registered are calculated. Two matched points, optimal and suboptimal, are found by KNN. Lowe's algorithm is adopted to compare whether the ratio of the two distances is less than a given threshold [35]. If so, the current matching relationship will be regarded as an acceptable matched point pair. Otherwise, the matching relationship will be eliminated.

2.4. Bi-Directional Fine Matching

The specific means of feature point detection and feature vector construction applied in ORB significantly contribute to processing speeds. Still, the algorithm's performances are poor in terms of accuracy. Hence, it is necessary to accomplish the fine matching of feature points through an appropriate strategy. However, the matching procedures of the most common registration methods are unidirectional, which only search matching objects for feature points in reference images. There are many errors and omissions in their matching results. Focused on these problems, two bi-directional matching strategies are proposed in this work, as depicted in Figure 3.

Strategy 1: The primary idea of the first bi-directional matching strategy is to remove wrongly matched points in initial matching results. As presented in Figure 3a, its specific procedure is as follows. Suppose that the feature points detected in an image to be registered are marked as $\{p_1, p_2, \ldots, p_n\}$, and the feature points detected in its reference image are marked as $\{q_1, q_2, \ldots, q_m\}$. After forward matching, a set of matched feature points in the two images is obtained and denoted as $PQ = \{p_iq_j, 1 \leq I \leq n \text{ and } 1 \leq j \leq m\}$. Similarly, after backward matching, a set of matched feature points in the two images can be represented by $QP = \{q_sp_t, 1 \leq s \leq n \text{ and } 1 \leq t \leq m\}$. Then, the backward matching results are traversed according to the forward matching result. If it is found that there are two identical pairs of matched points, i.e., p_iq_j and q_tp_s with $i = s$ and $j = t$, the point pair p_iq_j (or q_tp_s) can be regarded as a correct matching relationship and will be preserved. Otherwise, the matched point pairs in the backward matching result will be eliminated. Thereby, a new set of matched feature points $\widetilde{P}\widetilde{Q}$ can be built more reliably through this bi-directional matching strategy.

Strategy 2: The other bi-directional matching strategy is to compensate missed feature points into unidirectional results. The concrete realization is elaborated as follows. If there is a matched point pair q_sp_t ($1 \leq s \leq n$ and $1 \leq t \leq m$) in the backward matching result, but q_s and p_t do not appear in the forward matching result PQ, q_sp_t will be deemed as a matching pair missed in forward matching. Accordingly, the point pair p_tq_s will be added to the set PQ. Then, a new set $\widetilde{P}\widetilde{Q}$ containing more matched feature points is obtained until all the point pairs in the backward matching result are examined. The detail matching flow is elaborated in Figure 3b.
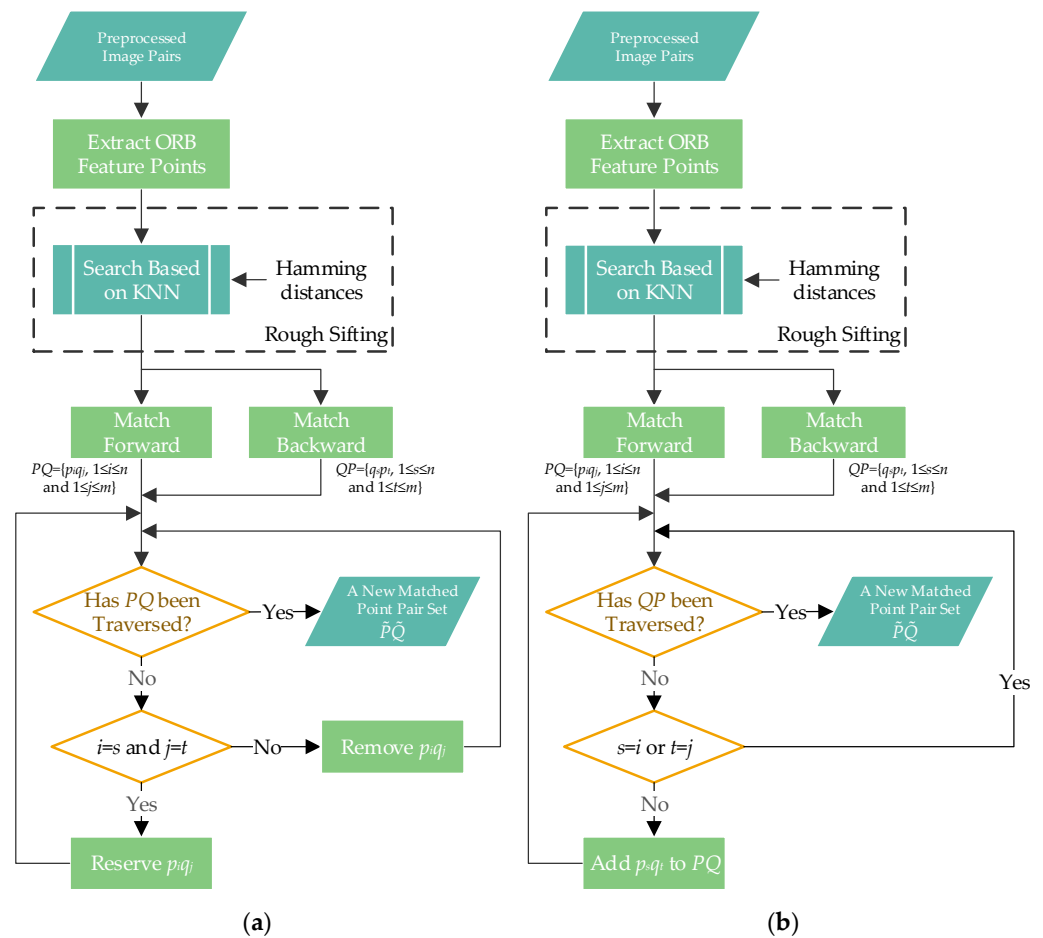
Figure 3. The flow chart of bi-directionally matching UAV images: (**a**) Strategy 1; (**b**) Strategy 2.

2.5. False Match Elimination and Transform Model Fitting

In this work, PROSAC is employed to remove mismatched feature points further from bi-directional rough matching results [36]. It is known that the classical RANSAC method is realized by random sampling of rough matching results to fit transform models. This method is random and heavily depends on pre-set iteration times for fitting models. Nevertheless, PROSAC is constructed by sampling an increasing set of optimally matched point pairs, which have been ranked based on a certain similarity measure. Sampling and fitting models based on top-ranked point pairs help enhance the success rates of obtaining correct models. Meanwhile, the randomness of the algorithm decreases obviously. Specifically, matched point pairs in $\widetilde{P}\widetilde{Q}$ are sorted from near to far according to Hamming distances of feature descriptors. Then, four point pairs are randomly chosen from the top m pairs in the reordered feature point-pair set, which can be used to fit a 3×3 homography matrix H, which can satisfy the following condition.

$$\begin{bmatrix} x\prime \\ y\prime \\ 1 \end{bmatrix} \sim H_{3\times 3} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{11}$$

where (x, y) and (x', y') correspond to point coordinates in the image to be registered and the reference image, respectively. According to the current parameters of the transform matrix H, the coordinates of all other points in the image to be registered are transformed by H into the coordinate system of the reference image. Then, these points are classified into inner or outlier points according to the differences between their projection coordinates and corresponding point coordinates in the reference image. These inner points are refined

matched point pairs, and H will be renewed in light of them. Repeatedly, the inner point set and H will be updated until the pre-set times of iterations are reached.

3. Results

The experimental hardware platform includes an Intel Core i5-4590K processor with a main frequency of 3.5 GHz and 16 GB RAM. The experimental software environment contains a 64-bit operating system Windows 10, a programming tool Visual Studio 2017, and an open-source library OpenCV 3.20. The comprehensive performances of the proposed registration algorithm are verified in the experiments by using UAV images of different sensors acquired from different scenes, including urban, road, building, farmland, and forest. Moreover, the improved ORB algorithm is compared with the existing popular algorithms, including SIFT [37], SURF [38], KAZE (a Japanese word) [39], AKAZE (Accelerated KAZE) [40], and its original version in terms of speed and accuracy.

3.1. Determining the Number of Feature Points

Since the number of feature points extracted by the ORB operator is optional, a comparative experiment is carried out to determine an appropriate number of feature points for optimal registration effects. Hence, the UAV images of five scenes are registered through the different numbers of feature points extracted by the original ORB algorithm. The relationship between the number of feature points and running time or Root Mean Square Error (*RMSE*) are separately plotted in Figure 4a,b. *RMSE* can be denoted as follows [2].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\|T(\boldsymbol{y}_i, \boldsymbol{\theta}) - \boldsymbol{x}_i\|\right)^2} \tag{12}$$

where x_i and y_i (for $\underline{i}$ = 1, 2, . . ., n) stand for one of the n matching point pairs from the image to be registered and the reference image. T is a transform model, and $\boldsymbol{\theta}$ is the model parameter vector. $\| \bullet$ is the Euclidean distance between two points. Generally, the smaller the value of RMSE, the higher the registration accuracy. From Figure 4a, it can be found that the running time of the algorithm is directly proportional to the number of feature points. Meanwhile, it can be seen from Figure 4b that there exist minimum RMSE values individually at 1000, 2000, and 2500, and most of the values of RMSE for the five senses reach the lowest when the number of feature points is 1000. Thereby, considering running time and registration accuracy, the number of feature points to be extracted is uniquely set at 1000 for different scenes in the subsequent registration experiments.
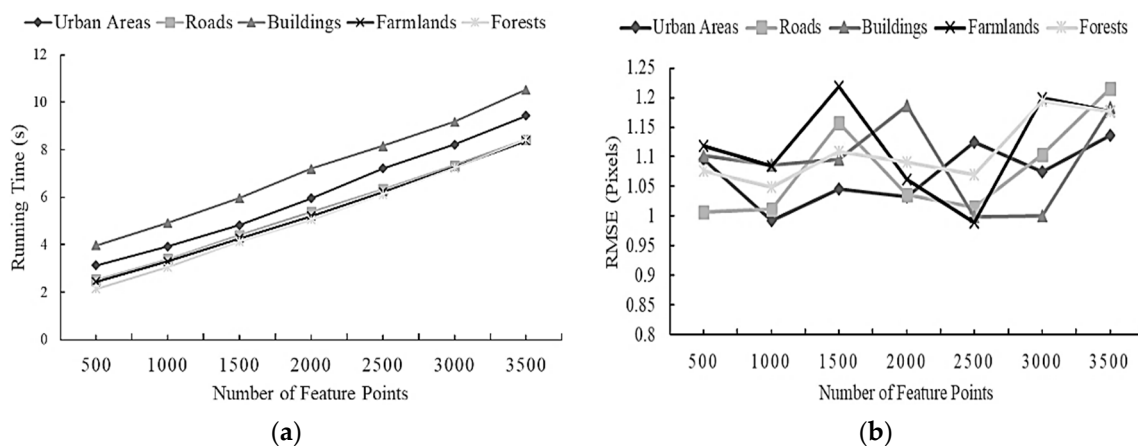

Figure 4. The relationship between the number of feature points and running time or RMSE (Root Mean Square Error) for five scenes: (**a**) the number of feature points vs. running time; (**b**) the number of feature points vs. RMSE.

3.2. Registration of Scene Images

The sources and parameters of UAV images used in the registration experiments are listed in Table 1. The urban and building scene images are, respectively, selected from the high-altitude and near-ground drone data of an ISPRS benchmark for multi-platform photogrammetry [41]. The image pairs of a road scene are the test images provided by Pix4Dmapper v.2.0 software. The farmland and forest scene images are self-collected. Particularly, the image pair of a farmland scene was acquired by the Parrot Sequoia multi-spectral camera with four bands, including green, red, red-edge, and near-infrared. Its red edge images are employed as blue components in the registration of this work. The two images of each scene image-pair were taken at different angles with an overlap. The image pair of the urban scene possesses many details and is quite complex, as shown in Figure 5. As mentioned before, 1000 pairs of feature points are extracted in the two images, and only 99 pairs of points are left after initial screening, as displayed in Figure 5c.

Table 1. The sources and parameters of UAV images used in the registration experiments.

Scenes	Sensors	Image Resolution	Plots	Shooting Time	Shooting Height (m)	Focal Length (mm)
urban	Hasselblad H3DII-50	2044 × 1533	Toronto, Canada	June 2014	500	80
road	Canon IXUS 220HS	2000 × 1500	Brig, Switzerland	September 2013	200	4
building	SONY NEX-7	3000 × 2000	Dortmund, Germany	June 2014	50	16
farmland	Parrot Sequoia	2404 × 1728	Dayi, China	September 2017	80	5
forest	ZENMUSE Z30	1920 × 1080	Wusu Foshan Forest Park, China	June 2019	152	10

Based on the result of rough matching, two bi-directional matching strategies are adopted in fine matching, as shown in Figure 6. There are 60 matched point pairs preserved after screening out 39 point pairs using the first bi-directional matching strategy. After processing by the second bidirectional matching strategy, 121 matching point pairs remain, including 22 supplemental point pairs.

Then, the elimination results of false matches using the PROASC algorithm are shown in Figure 7, and the final registration results are presented in Figure 8. Since some pixel points in transformed images cannot be assigned directly to new grid coordinates, bilinear interpolation is utilized to generate stitched images [42]. For the result of the first bi-directional matching strategy, 23 pairs of matched point pairs remain, and 37 pairs of matching point pairs are removed by the PROSAC algorithm. In this work, the running time and RMSE of the registration algorithm are chosen as the main indicators in performance comparisons. The running time of the registration process is 3.922 s, and the RMSE is 0.9926 pixels. The obtained transform matrix is as follows.

$$H = \begin{bmatrix} 1.000727295875549 & 0.000115774389996 & -675.5347900390625 \\ 0.000406843028031 & 1.000958800315857 & 118.6014556884766 \\ 0.000000566664652 & 0.000000197085213 & 1 \end{bmatrix} \tag{13}$$

For the result of the second bi-directional matching strategy, 45 pairs of matching point pairs are left, and 76 pairs of matching point pairs are removed by the PROSAC algorithm. The registration time is 4.031 s, and the RMSE is 1.0750 pixels. The corresponding perspective matrix is denoted as

$$H = \begin{bmatrix} 1.000828385353088 & 0.000052663752680 & -675.6309814453125 \\ 0.000271332741249 & 1.001049280166626 & 118.7060470581055 \\ 0.000000538613107 & 0.000000276266348 & 1 \end{bmatrix} \tag{14}$$
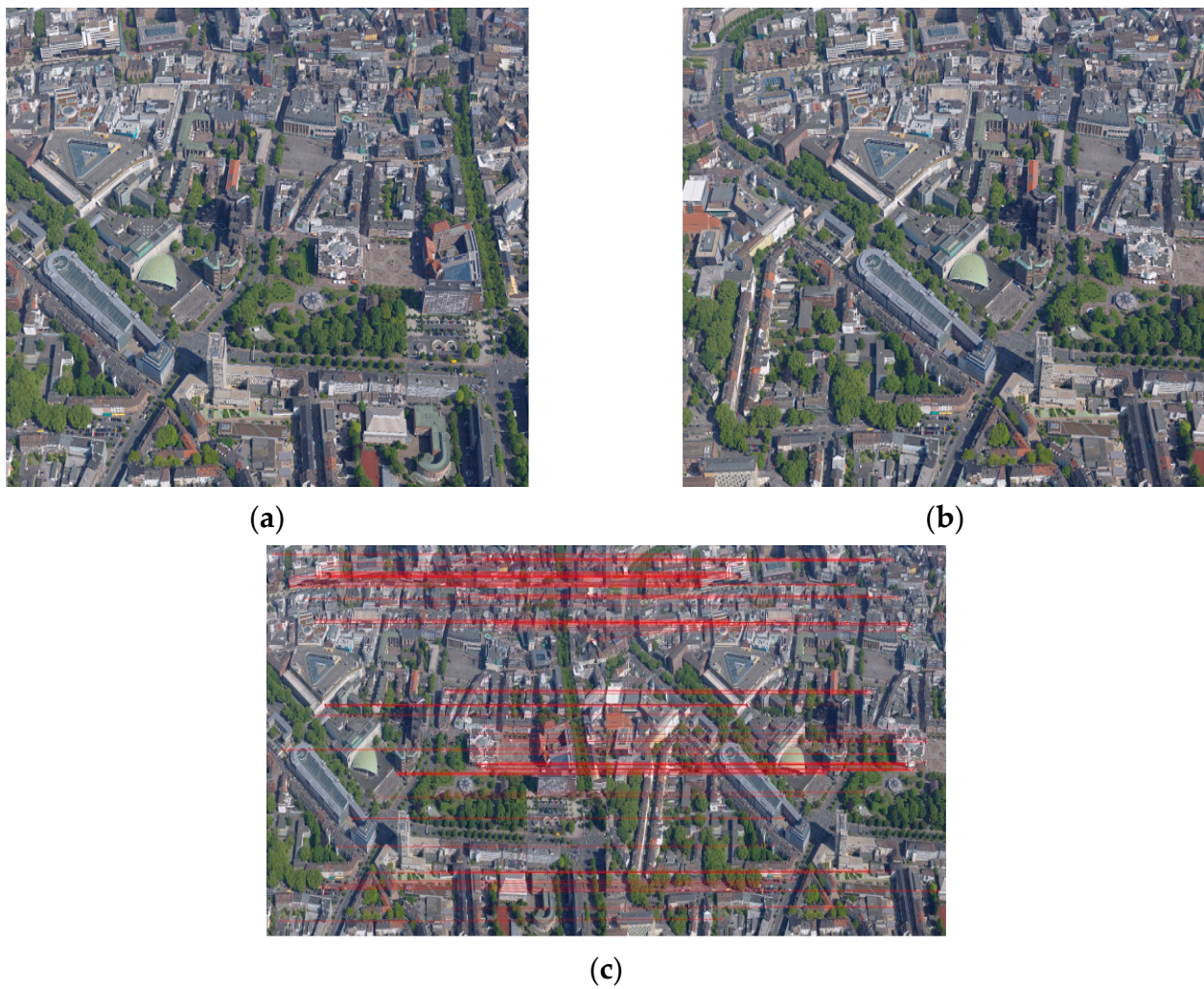
(a)

(b)

(c)

Figure 5. A UAV image pair from an urban scene and the feature points after the initial screening: (**a**) the image to be registered; (**b**) the reference image; (**c**) the image pair after the rough screening. The corresponding points are connected by red lines.
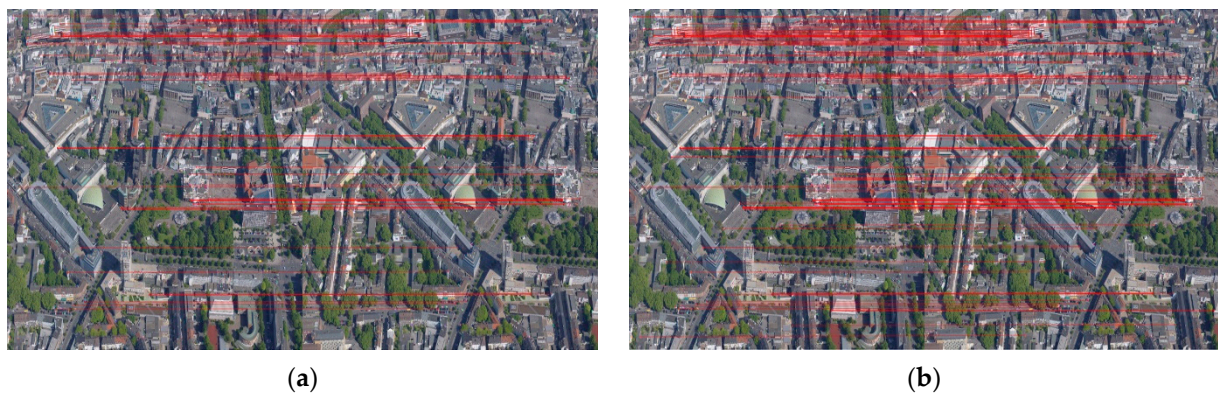

(a)

(b)

Figure 6. The urban image pairs after bi-directional matching according to two strategies: (**a**) Strategy 1; (**b**) Strategy 2. The corresponding points are connected by red lines.
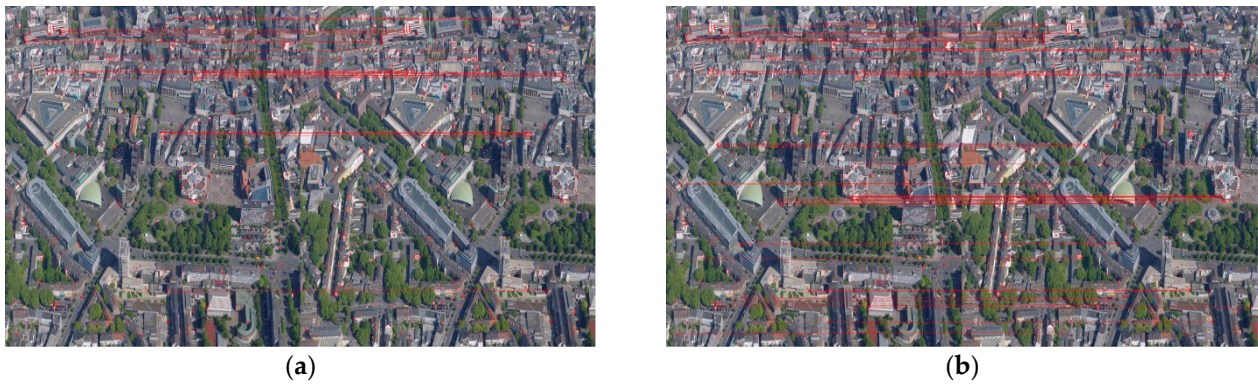
(**a**) (**b**)

Figure 7. The results of bi-directional matching after eliminating false matches: (**a**) Strategy 1; (**b**) Strategy 2. The corresponding points are connected by red lines.
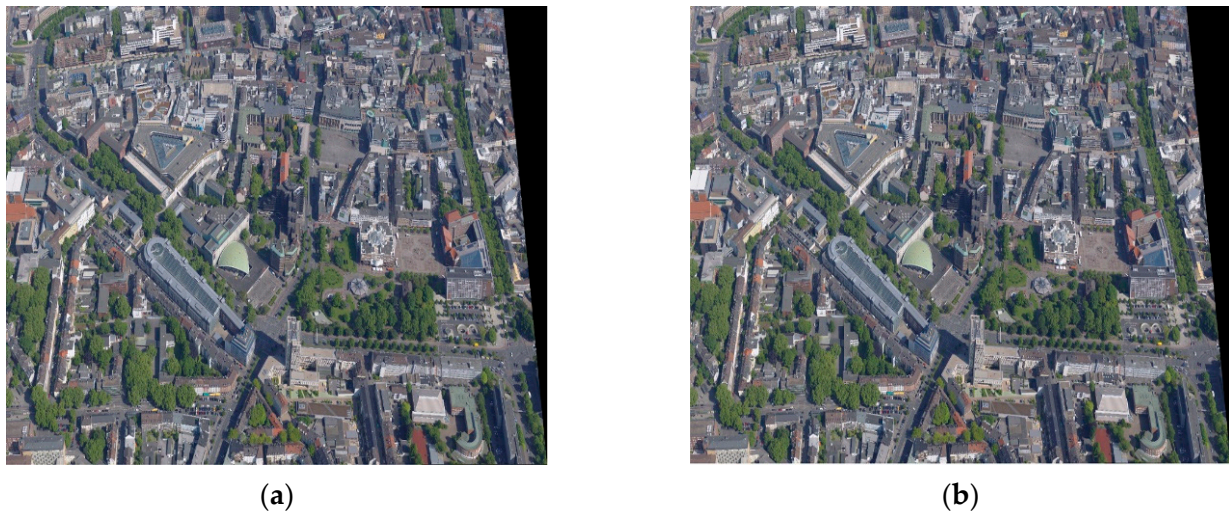

(**a**) (**b**)

Figure 8. The stitched image of the urban scene by the proposed registration algorithms: (**a**) Strategy 1; (**b**) Strategy 2.

The registration accuracies of different registration algorithms for different scenes are listed in Table 2. For the five scene images, it can be seen that regarding accuracies, the proposed registration algorithms can provide obvious improvements on the original ORB algorithm, and the method based on the first bi-directional matching strategy exceeds all the other algorithms. For the scene images of urban areas, buildings, and farmlands, our algorithm even can achieve sub-pixel registration accuracies. In addition, it is revealed that the accuracy improvements of the proposed methods are related to the scenes of the experimental images. If there are rich details in images such as those found in urban, road, and building images, the accuracies can increase by more than 20.0% compared to the original ORB algorithm. If the images have fewer details, such as farmland and forest images, the accuracy enhancements are not outstanding, but still about 10.0%. The reason may be that the more details in images, the greater the differences in extracted feature points. Furthermore, the qualities of matched feature point pairs can be improved greatly after coarse and fine screening so that registration accuracies increase.

Table 2. RMSE comparison of registration accuracy of different algorithms (pixels).

Scenes \ Methods	SIFT	SURF	KAZE	AKAZE	ORB	Ours (Strategy 1)	Ours (Strategy 2)
urban	1.2322	1.1242	1.1846	1.0206	1.3282	**0.9926**	1.0750
road	1.1805	1.2670	1.2945	1.1546	1.3530	**1.0124**	1.1757
building	1.1758	1.2918	1.2173	1.1163	1.3273	**0.9828**	1.0471
farmland	1.3735	1.3905	1.2668	1.1606	1.2005	1.0853	**0.9909**
forest	1.2692	1.3344	1.2287	1.2327	1.2871	**1.0490**	1.0776

The best results for different scene are bolded.

The time consumption comparison of different registration algorithms is presented in Table 2. As can be seen, the ORB algorithm has obvious advantages in running time over the other five algorithms. The reason may be that the high resolutions of these experimental UAV images result in more feature points being detected by the other five algorithms. Consequently, it takes a lot of time for these methods, especially for the KAZE algorithm, to detect and match feature points. However, it seems rational that the matching accuracy of the ORB algorithm is the lowest in Table 3 because its performance is closely related to the number of feature points. Our proposed registration methods cannot only inherit the advantage in speeds from the ORB algorithm but also achieve excellent registration qualities by combining the bi-directional matching strategies and a false-match elimination method. Hence, it can be drawn that our proposed registration algorithm realizes accuracy enhancement at the small cost of processing speeds.

Table 3. Time consumption comparison of different registration algorithms (s).

Scenes \ Methods	SIFT	SURF	KAZE	AKAZE	ORB	Ours (Strategy 1)	Ours (Strategy 2)
urban	125.254	64.023	128.408	101.053	2.252	3.922	4.031
road	137.025	86.799	180.031	144.146	2.324	3.483	3.572
building	163.348	139.725	194.028	117.054	3.457	4.948	4.992
farmland	50.899	46.652	72.712	35.768	1.414	3.562	3.775
forest	93.889	42.794	84.774	103.447	1.727	3.315	3.295

In addition, the numbers of matched point pairs in the five scene images after different process stages are illustrated in Figure 9. It is known that there are abundant detail features in the UAV images of urban areas, buildings, and roads. In contrast, the contents of the UAV images of farmlands and forests are relatively simple, and their feature textures are not enough. In general, a sufficient number of feature points need to be extracted by matching methods effectively for fitting transformation model parameters. For the scene images of urban areas, buildings, and roads, a relatively large number of representative feature points can be detected. Although it benefits the increases in registration accuracies, it also spends more time in additional calculations. For the images of farmlands and forests, fewer and lower representative feature points are detected, which leads to lower registration accuracies but less calculation time.

The results of bi-directional matching for the other four scene image pairs are in Figure 10, and the final stitched images are displayed in Figure 11. It can be observed from Figure 11a that in the road scene, there are slight mismatches for one road in the mosaicking result of the first bi-directional matching strategy. The reason may be that the less matched point is extracted by this strategy. In Figure 11b, there are some obvious boundary effects in the body of a transmission tower, which may be due to its incomplete appearance at the edge of the reference image. From Figure 11c,d, it can be thought that though there are no significant edge features such as points and lines in the scenes of farmland and forest, the proposed fast registration provides good results for these image pairs. It can be noted that only a visible mismatch for one road exists in the result of the first bi-directional matching strategy for the farmland scene. This may be owing to its insufficient matched

point pairs, as shown in Figure 10c. However, some areas for road and farmland scenes appear blurred owing to the quality of input images. Specifically, the red edge images of the farmland scene regarded as blue components in registration are inherently not clear enough. In addition, imaging conditions, such as shooting heights, sensor characteristics, atmospheric visibility and wind speed, light intensity, and the stability of drone platforms, may all affect the quality of drone images.
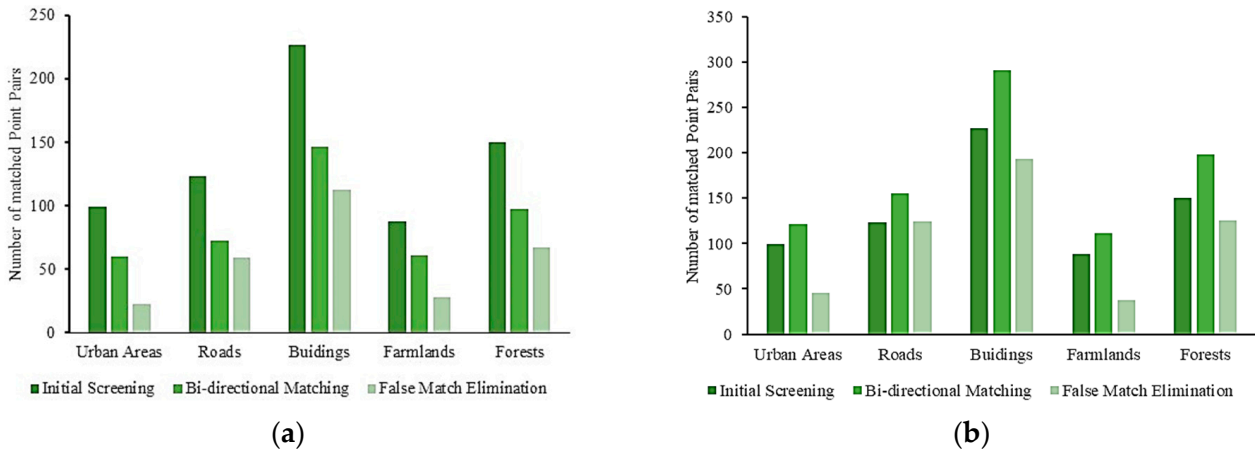

(**a**)

(**b**)

Figure 9. The numbers of matched point pairs in registration processes: (**a**) the first bi-directional strategy; (**b**) the second bi-directional strategy.
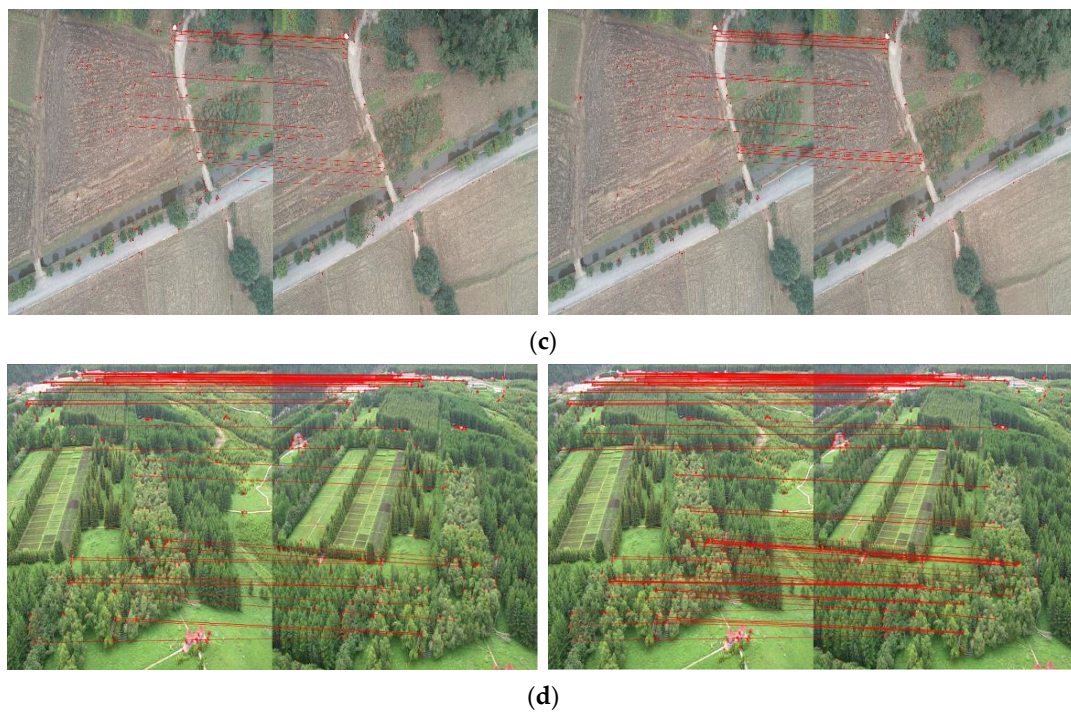

(**a**)

(**b**)

Figure 10. *Cont.*

(**c**)

(**d**)
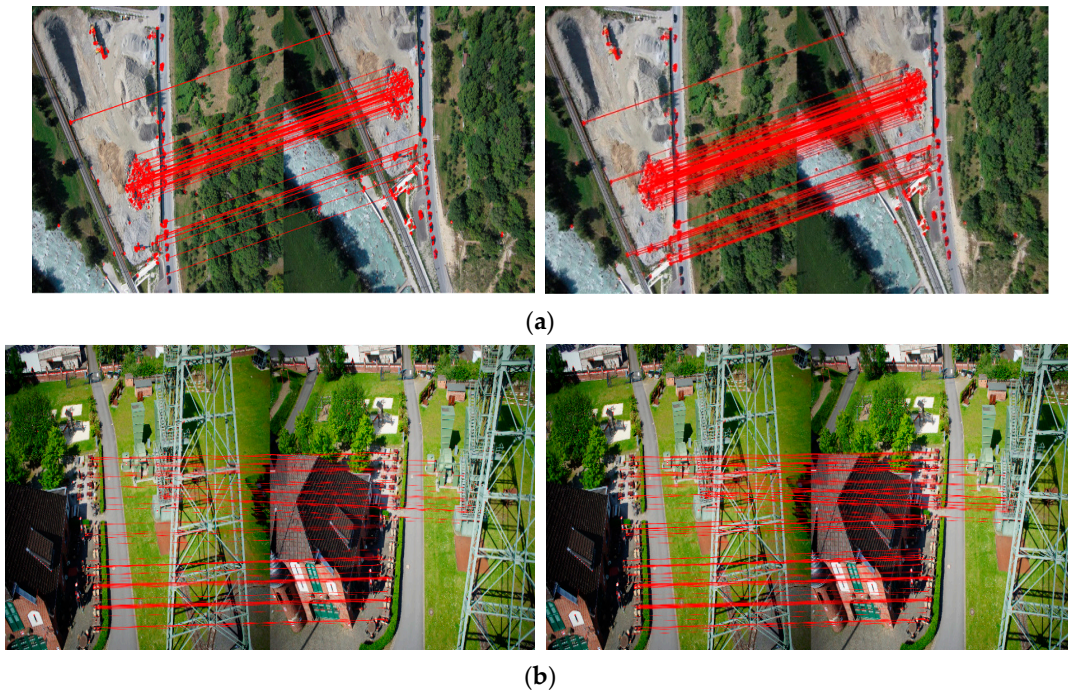
Figure 10. The different image pairs after bi-directional matching according to two strategies: (**a**) roads; (**b**) buildings; (**c**) farmlands; (**d**) forests. The corresponding points are connected by red lines.
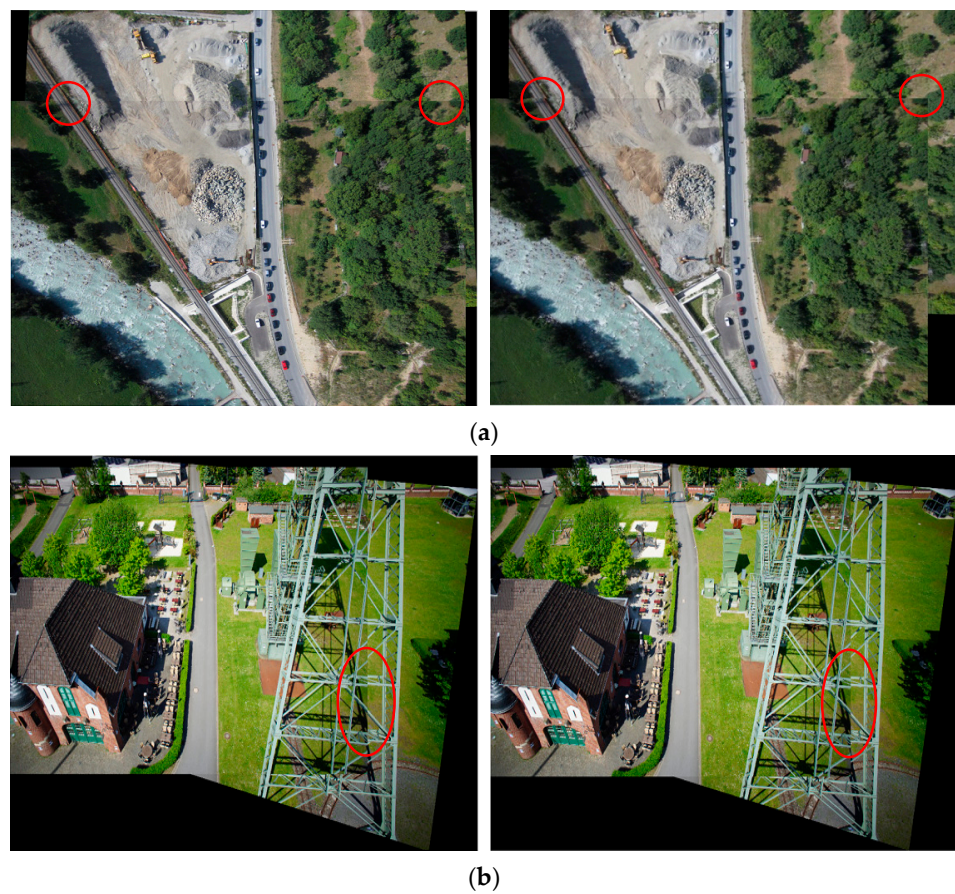

(**a**)
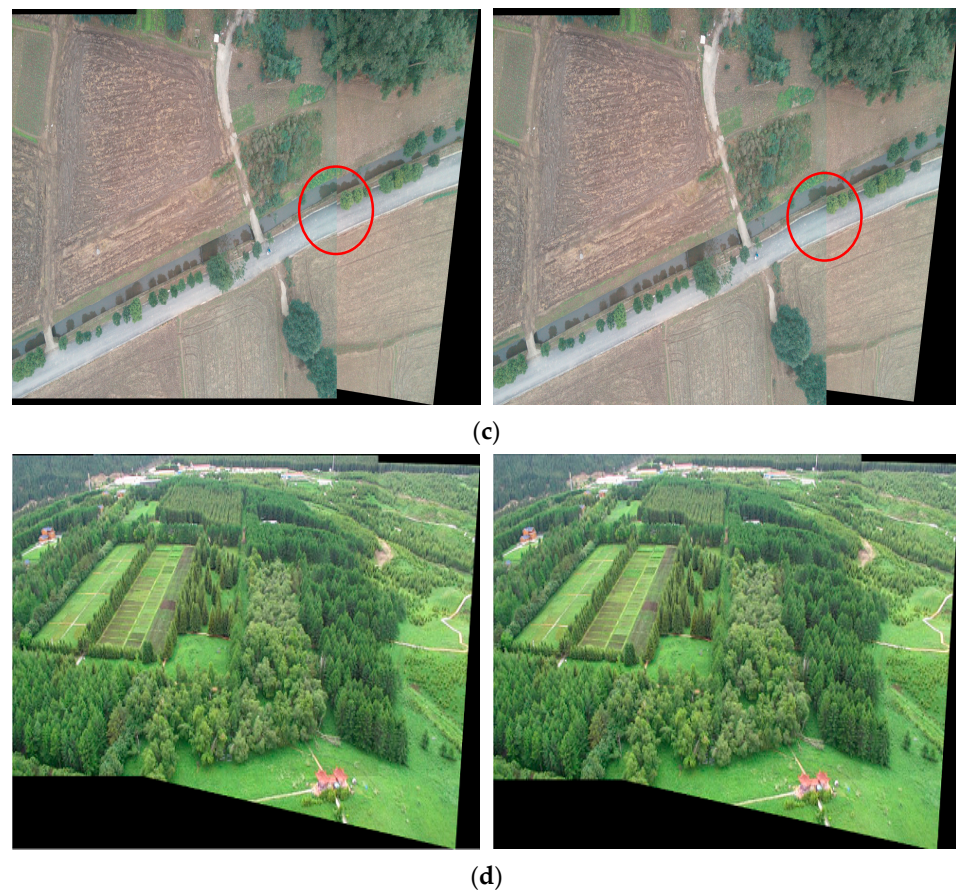

(**b**)

Figure 11. *Cont.*

(**c**)

(**d**)

Figure 11. The stitched images of four scenes by the proposed registration algorithms. The left one is obtained according to the first bi-directional matching strategy, and the right one is generated by using the second bi-directional matching strategy: (**a**) roads; (**b**) buildings; (**c**) farmlands; (**d**) forests. Some obvious mismatches are marked by red circles.

4. Conclusions

Since the ORB algorithm is not able to balance accuracy and time in the registration of UAV-visible images, in this work, a fast automatic registration method for UAV images is proposed by combining the ORB operator with bi-directional matching strategies. The KNN-based search method and a similarity measurement are assembled in initial matching. Two bidirectional matching strategies are designed for fine matching, and the PROSAC algorithm is employed to remove mismatches. In order to verify its performance, the proposed algorithm is compared with the existing fast algorithm for the registration of UAV images from different sensors and scenes. From the experimental results, it can be thought that the improvements of this work are effective. The proposed registration algorithm can enhance matching accuracy for UAV images without requiring any training or prior knowledge while maintaining the high registration speeds of the original ORB algorithm. The maximum accuracy improvement for the experimental images reaches 25.95% for the building scene. In terms of processing time and accuracy, the bi-directional matching method of reducing matched point pairs slightly outperforms the method of increasing matched point pairs. In addition, the proposed registration algorithm needs no training or prior knowledge with low complexities, which is quite suitable for onboard realization.

Author Contributions: Conceptualization and methodology X.L. and X.W. (Xiao Wang); software and validation, X.W. (Xiao Wang) and Z.W.; formal analysis, X.L. and Z.W.; resources, W.Z.; data curation and writing—original draft preparation, X.L. and P.L.; writing—review and editing, X.L. and X.W. (Xufeng Wei); visualization and project administration, W.Z. and Y.J.; funding acquisition, W.Z. All authors have read and agreed to the published version of the manuscript.

References

1. Etezadifar, P.; Farsi, H. A new sample consensus based on sparse coding for improved matching of SIFT features on remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 5254–5263. [CrossRef]
2. Luo, X.; Lai, G.; Wang, X.; Jin, Y.; He, X.; Xu, W.; Hou, W. UAV remote sensing image automatic registration based on deep residual features. *Remote Sens.* **2021**, *13*, 3605. [CrossRef]
3. Meng, S. A robust approach for bridge crack image mosaic based on SURF algorithm. *Comput. Sci. Appl.* **2019**, *9*, 375–383. [CrossRef]
4. Goel, P.; Dayal, P.; Patra, T.K.; Gupta, C. Automatic co-registration of aerial and satellite imagery using optimized minimum distance thresholds for matching. In Proceedings of the 2022 International Mobile and Embedded Technology Conference (MECON), Noida, India, 10–11 March 2022. [CrossRef]
5. Gong, X.; Liu, Y.; Yang, Y. Robust stepwise correspondence refinement for low-altitude remote sensing image registration. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 1736–1740. [CrossRef]
6. Liu, Y.; Gong, X.; Yang, Y. A multilayer fusion network with rotation-invariant and dynamic feature representation for multiview low-altitude image registration. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 1019–1023. [CrossRef]
7. Liu, Y.; Liu, Y.; Wang, J. RFM-GAN: Robust feature matching with GAN-based neighborhood representation for agricultural remote sensing image registration. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 1–5. [CrossRef]
8. Silva, L.H.F.P.; Junior, J.D.D.; Mari, J.F.; Escarpinati, M.C.; Backes, A.R. Non-linear co-registration in UAVs' images using deep learning. In Proceedings of the 35th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Natal, Brazil, 24–27 October 2022. [CrossRef]
9. Sun, Y.; Qi, H.; Wang, C.; Tao, L. Image registration method based on generative adversarial networks. In Proceedings of the 2020 Eighth International Conference on Advanced Cloud and Big Data (CBD), Taiyuan, China, 5–6 December 2020. [CrossRef]
10. Hou, X.; Gao, Q.; Wang, R.; Luo, X. Satellite-borne optical remote sensing image registration based on point features. *Sensors* **2021**, *21*, 2695. [CrossRef] [PubMed]
11. Qiao, M.; Liang, X.; Chen, M. Improved SIFT algorithm based on image filtering. In Proceedings of the Journal of Physics: Conference Series, Sanya, China, 29–31 January 2021. [CrossRef]
12. Zhao, M.; Chen, H.; Song, T.; Deng, S. Research on image matching based on improved RANSAC-SIFT algorithm. In Proceedings of the 16th International Conference on Optical Communications and Networks (ICOCN), Wuzhen, China, 7–10 August 2017. [CrossRef]
13. Li, X.; Luo, X.; Wu, Y.; Li, Z.; Xu, W. Research on stereo matching for satellite generalized image pair based on improved SURF and RFM. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Online, 26 September–2 October 2020. [CrossRef]
14. Wei, W.; Cheng, A.; Zhao, Y.; Zhang, G. Image registration algorithm based on superpixel segmentation and SURF feature points. In Proceedings of the 8th International Conference on Image Processing Theory, Tools and Applications (IPTA), Xi'an, China, 7–10 November 2018. [CrossRef]
15. Li, X.; Zhu, J.; Ruan, Y. Vehicle seat detection based on improved RANSAC-SURF algorithm. *Int. J. Pattern Recognit. Artif. Intell.* **2021**, *35*, 2155004. [CrossRef]
16. Al Tamimi, A.; Qasaimeh, A.; Qaddoum, K. Offline signature recognition system using oriented FAST and rotated BRIEF. *Int. J. Electr. Comput. Eng. (IJECE)* **2021**, *11*, 4095–4103. [CrossRef]
17. Zhai, Y.; Liu, S.; Guo, X.W.; He, P.; Zhang, Z.H. Lunar image matching based on FAST features with adaptive threshold. In Proceedings of the 7th International Conference on Communications, Signal Processing, and Systems (CSPS), Dalian, China, 14–16 July 2018. [CrossRef]
18. Calonder, M.; Lepetit, V.; Strecha, C.; Fua, P. Brief: Binary robust independent elementary features. In Proceedings of the Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, 5–11 September 2010. [CrossRef]
19. Yang, Z.; Qin, X. Image stitching technology based on SIFT, FLANN, and RPOSAC algorithms. In Proceedings of the 2023 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL), Zhuhai, China, 12–14 May 2023. [CrossRef]

20. Saravanan, C. Color image to grayscale image conversion. In Proceedings of the 2010 Second International Conference on Computer Engineering and Applications (ICCEA), Bali, Indonesia, 26–29 March 2010. [CrossRef]

21. Li, S.; Wang, Z.; Zhu, Q. A research of ORB feature matching algorithm based on fusion descriptor. In Proceedings of the 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 12–14 June 2020. [CrossRef]

22. Liu, Y.; Wang, Y. A new ORB feature matching image algorithm based on non-maximum suppression method and Retinal sampling model. In Proceedings of the 2021 International Conference on Control, Automation and Information Sciences (ICCAIS), Xi'an, China, 14–17 October 2021. [CrossRef]

23. Sun, C.; Qiao, N.; Sun, J. Robust feature matching based on adaptive ORB for vision-based robot navigation. In Proceedings of the 2021 36th Youth Academic Annual Conference of Chinese Association of Automation (YAC), Nanchang, China, 28–30 May 2021. [CrossRef]

24. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the 2011 IEEE International Conference on Computer Vision (ICCV), Barcelona, Spain, 6–13 November 2011. [CrossRef]

25. Ihmeida, M.; Wei, H. Image registration techniques and applications: Comparative study on remote sensing imagery. In Proceedings of the 2021 14th International Conference on Developments in eSystems Engineering (DeSE), Sharjah, United Arab Emirates, 7–10 December 2021. [CrossRef]

26. Setyadi, R.; Ranggadara, I. Augmented reality using features accelerated segment test for property catalogue. *Telecommun. Comput. Electron. Control* **2020**, *18*, 140–147. [CrossRef]

27. Heo, H.; Lee, J.; Lee, K.; Lee, C. FPGA-based implementation of FAST and BRIEF algorithm for object recognition. In Proceedings of the 2013 IEEE International Conference of IEEE Region 10 (TENCON 2013), Xi'an, China, 22–25 October 2013. [CrossRef]

28. Li, Q.; Han, P.; Liu, D. Anti-fake legend registration algorithm based on improved ORB. In Proceedings of the 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 11–13 December 2020. [CrossRef]

29. Taranco, R.; Arnau, J.; Gonzalez, A. LOCATOR: Low-power ORB accelerator for autonomous cars. *J. Parallel Distrib. Comput.* **2023**, *174*, 32–45. [CrossRef]

30. Li, J.; Lin, S.; Yu, K.; Guo, G. Quantum K-nearest neighbor classification algorithm based on Hamming distance. *Quantum Inf. Process.* **2022**, *21*, 18. [CrossRef]

31. Mitsuya, S.; Nakashima, Y.; Inenaga, S.; Bannai, H.; Takeda, M. Compressed communication complexity of hamming distance. *Algorithms* **2021**, *14*, 116. [CrossRef]

32. Qin, M. Hamming-distance-based binary representation of numbers. In Proceedings of the 2018 Information Theory and Applications Workshop (ITA), San Diego, CA, USA, 11–16 February 2018. [CrossRef]

33. Cunningham, P.; Delany, S.J. k-Nearest Neighbour Classifiers–A Tutorial. *ACM Comput. Surv.* **2021**, *54*, 1–25. [CrossRef]

34. de Lima, R.; Cabrera-Ponce, A.A.; Martinez-Carranza, J. Parallel hashing-based matching for real-time aerial image mosaicing. *J. Real-Time Image Process.* **2021**, *18*, 143–156. [CrossRef]

35. Wang, X.; Cao, W.; Yao, C.; Yin, H. Feature matching algorithm based on SURF and Lowes algorithm. In Proceedings of the 39th Chinese Control Conference (CCC), Shenyang, China, 27–29 July 2020. [CrossRef]

36. Ma, S.; Ma, X.; You, H.; Tang, T.; Wang, J.; Wang, M. SC-PROSAC: An improved progressive sample consensus algorithm based on spectral clustering. In Proceedings of the 2021 3rd International Conference on Robotics and Computer Vision (ICRCV), Beijing, China, 6–8 August 2021. [CrossRef]

37. Lopez, A.; Jurado, J.M.; Ogayar, C.J.; Feito, F.R. A framework for registering UAV-based imagery for crop-tracking in precision agriculture. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *97*, 102274. [CrossRef]

38. Jhan, J.; Rau, J. A generalized tool for accurate and efficient image registration of UAV multi-lens multispectral cameras by N-SURF matching. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 6353–6362. [CrossRef]

39. Kerkech, M.; Hafiane, A.; Canals, R. Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach. *Comput. Electron. Agric.* **2020**, *174*, 105446. [CrossRef]

40. Yan, J.; Wang, Z.; Wang, S. Real-time tracking of deformable objects based on combined matching-and-tracking. *J. Electron. Imaging* **2016**, *25*, 023019. [CrossRef]

41. Nex, F.; Gerke, M.; Remondino, F.; Przybilla, H.-J.; Bäumker, M.; Zurhorst, A. ISPRS benchmark for multiplatform photogrammetry. In Proceedings of the ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Munich, Germany, 25–27 March 2015. [CrossRef]

42. Ye, H.; Su, K.; Huang, S. Image enhancement method based on bilinear interpolating and wavelet transform. In Proceedings of the 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 12–14 March 2021. [CrossRef]