sensors

MDPI

Article

Image Stitching Based on Color Difference and KAZE with a Fast Guided Filter

Chong Zhang, Dejiang Wang * and He Sun

Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences,
Changchun 130033, China; chongzhang17@mails.jlu.edu.cn (C.Z.)
* Correspondence: wangdj04@ciomp.ac.cn

Abstract: Image stitching is of great importance for multiple fields such as moving object detection and tracking, ground reconnaissance and augmented reality. To ameliorate the stitching effect and alleviate the mismatch rate, an effective image stitching algorithm based on color difference and an improved KAZE with a fast guided filter is proposed. Firstly, the fast guided filter is introduced to reduce the mismatch rate before feature matching. Secondly, the KAZE algorithm based on improved random sample consensus is used for feature matching. Then, the color difference and brightness difference of the overlapping area are calculated to make an overall adjustment to the original images so as to improve the nonuniformity of the splicing result. Finally, the warped images with color difference compensation are fused to obtain the stitched image. The proposed method is evaluated by both visual effect mapping and quantitative values. In addition, the proposed algorithm is compared with other current popular stitching algorithms. The results show that the proposed algorithm is superior to other algorithms in terms of the quantity of feature point pairs, the matching accuracy, the root mean square error and the mean absolute error.

Keywords: image stitching; KAZE; RANSAC; fast guided filter

1. Introduction

Currently, image stitching is still a topical research issue and is widely used in computer vision, unmanned aerial vehicle (UAV) reconnaissance and other fields [1,2]. Image stitching consists of synthesizing different types of images captured by different imaging devices at different shooting positions into an image with a larger field of vision [3,4]. On the one hand, image stitching can synthesize panoramic and ultra-wide-view images, so that ordinary cameras can achieve grand scene shooting. On the other hand, image stitching can synthesize fragmented images into a complete image. The stitching technology can be applied to combine medical images, scientific microscope fragments or local images from seabed exploration into a complete image. In addition, image mosaic is also a basic technology in scene rendering methods, which uses panoramic images instead of 3D scenes to model and draw. However, image stitching still encounters many challenges at present, such as parallax caused by viewpoint change, ghosting, distortion, detail distortion and image unevenness [5–7].

Up to now, many image stitching methods have been proposed by researchers. Li et al. [8] proposed a model of appearance and motion variation based on the traditional speeded up robust features (SURF) algorithm, which mainly contains Hessian matrix construction, Hessian matrix determinant approximation calculations and non-maximal suppression determination feature points in its key steps. In order to improve the matching accuracy and robustness, Liu et al. [9] introduced an improved random sample consensus (RANSAC) feature image matching method based on SURF. First of all, the SURF method is used to detect and extract image features, and the fast library of the approximate nearest neighbor-based matcher method is used to perform initial matching on image feature points.

check for updates

Citation: Zhang, C.; Wang, D.; Sun, H. Image Stitching Based on Color Difference and KAZE with a Fast Guided Filter. *Sensors* **2023**, *23*, 4583. https://doi.org/10.3390/s23104583

Received: 25 March 2023
Revised: 2 May 2023
Accepted: 5 May 2023
Published: 9 May 2023

The RANSAC algorithm is improved to increase the probability of correct matching points. In 2017, Guan et al. [10] presented an interest point detector and binary feature descriptor for spherical images. Inspired by the Binary Robust Invariant Scalable Keypoints (BRISKs), they adapted the method to operate on spherical images. All of the processing is intrinsic to the sphere and avoids the distortion inherent in storing and indexing spherical images in a 2D representation. Liu et al. [11] developed the BRISK_D algorithm, which effectively combines features from the Accelerated Segment Test (FAST) and BRISK methods. The keypoints are detected by the FAST algorithm and the location of the keypoint is refined in scale and space. The scale factor of the keypoint is directly computed with the depth information of the image. Zhang et al. [12] proposed a screening method based on binary mutual information for the mismatch problem. The feature points extracted by the ORB algorithm are distributed in a color change area. Then, the new feature points are obtained by internal points. In this way, feature points can be eliminated and the best transformation matrix can be obtained by using an iterative method. In 2020, Ordóñez et al. [13] introduced a registration method for hyperspectral remote sensing images based on MSER, which effectively utilized the information contained in different spectral bands. Elgamal et al. [14] proposed an improved Harris hawks optimization by utilizing elite opposite-based learning and proposing a new search mechanism. This method can avoid falling into local optimum, improve the calculation accuracy and accelerate the convergence speed. Debnath et al. [15] utilized Min Eigen feature extraction based on the Shi-Tomasi corner detection method which detects interest points to identify image forgery.

With the continuous development of machine learning, many scholars have applied it to the field of image stitching. However, learning-based image stitching solutions are rarely studied due to the lack of labeled data, making the supervised methods unreliable. To address this limitation, Lang Nie et al. [16] proposed an unsupervised deep image stitching framework consisting of two stages: unsupervised coarse image alignment and unsupervised image reconstruction. The stitching effect of this method is more obvious for images with few features or low resolution. However, stitching algorithms based on machine learning require extensive training, have a high resource consumption and are highly time consuming.

Alcantarilla et al. [17] introduced the KAZE feature in 2012. Their results revealed that it performs better than other feature-based detection methods in terms of detection and description. Pourfard et al. [18] proposed an improved version of the KAZE algorithm with the accelerated robust feature (SURF) descriptor for SAR image registration. Discrete random second-order nonlinear partial differential equations (PDEs) are used to model the edge structure of SAR images. The KAZE algorithm uses nonlinear diffusion filtering to build up the scale levels of the SIFT descriptor. It preserves edges while smoothing the image and reduces speckle noise.

For the purpose of alleviating the sensitivity of KAZE and improving the stitching effect, we proposed image stitching based on color difference and KAZE with a fast guided filter as an efficient stitching method. The contributions of this paper are as follows: firstly, we introduce a fast guided filter into the KAZE algorithm to effectively reduce the mismatch information and improve the matching efficiency. Secondly, we use the improved RANSAC algorithm to increase the probability of correct matching points sampled and effectively eliminate the wrong matching point pairs. Thirdly, we introduce color difference and brightness difference to compensate for the whole image when fusing and stitching images. This method cannot just only effectively eliminate seams, but also results in perfectly uniform color and luminance.

The remainder of the paper is organized as follows. Section 2 depicts the fast guided filter and Additive Operator Splitting (AOS) algorithm briefly. Section 3 details the proposed method in this paper. Section 4 presents the experimental results and assessments. Finally, the conclusions and outstanding issues are listed in Section 5.

2. Related Work

In this section, we introduce the fast guided filter, nonlinear diffusion and the AOS algorithm briefly.

2.1. Fast Guided Filter

A guided filter is an adaptive weight filter which can restore the boundaries while smoothing images via a guide map [19]. Assume that the input image, the output image and the guidance map are expressed as p, q and I, respectively. A guided filter driven by a local linear model can be defined as shown in Equation (1):

$$q_i = a_k I_i + b_k, \forall i \in \omega_k, \tag{1}$$

where i is the index of pixels and ω expresses the selected local square window. a_k and b_k are two constants in the window ω. Aiming at minimizing the reconstruction error between the input image and the output result, optimal values of a_k and b_k can be acquired by Equations (2) and (3).

$$a_k = \frac{\frac{1}{|\omega|}\sum_{i \in \omega_k} I_i p_i - \mu_k \overline{p}_k}{\sigma_k^2 + \varepsilon}, \tag{2}$$

$$b_k = \overline{p}_k - a_k \mu_k, \tag{3}$$

where μ_k and σ_k are the mean and variance of the guided image I in the window k, respectively, and ε is a regularization parameter controlling the degree of smoothness. The values of a_k and b_k will determine the weight of gradient information and smoothing information. It is obvious that b_k is approximately equal to the mean value of the pixel points when the value of a is small.

To reduce the time complexity, He et al. [20] proposed a fast guided filter on this basis. The filtering output is computed by Equation (4):

$$q_i{}' = \overline{a}_i I_i + \overline{b}_i, \tag{4}$$

where $\overline{a}_i$ and $\overline{b}_i$ are the averages of a and b, respectively, on the same centered window, which can be calculated by Equations (2) and (3), respectively. Firstly, the input image and the guided image are downsampled by scale s. Secondly, the output image q' can be calculated by Equation (4). Finally, q' is upsampled s times to obtain an output image with the same size as the original image. The time complexity of the method becomes $O(N/s^2)$.

In this paper, a fast guided filter is applied to preprocessing before feature matching. The fast guided filter not only removes mismatch point pairs effectively, but also protects the texture information of the image preferably. The specific method is shown in Section 3.

2.2. Nonlinear Diffusion and the AOS Algorithm

The nonlinear diffusion method describes the luminance of an image with the increase in scale as the divergence of a flow function that controls the diffusion process [21]. This method is normally described by nonlinear partial differential equations (PDEs) owing to the nonlinear properties of the differential equations involved in diffusing the luminance of the image to the nonlinear scale space [22]. The classical nonlinear diffusion formula can be described by Equation (5):

$$\frac{\partial L}{\partial t} = \mathrm{div}(c(x, y, t) \cdot \nabla L), \tag{5}$$

where div and ∇ represent the divergence and gradient operations, respectively.

Li et al. [23] introduced nonlinear diffusion filtering and its implementation in detail. Nonlinear diffusion is a good smoothing method, which can reduce noise while maintaining the peak shape. In order to reduce the calculation loss at the local edges, more smooth regions are selected instead of boundaries. The conduction function can be defined by Equation (6):

$$c(x,y,t) = g(|\nabla L_\sigma(x,y,t)|), \tag{6}$$

where ∇L_σ indicates the gradient of the Gaussian smoothing version of the original image. Since we chose to smooth both sides of the edge, the conduction function can be defined as follows:

$$g = \begin{cases} 1 & , |\nabla L_\sigma|_2 = 0 \\ 1 - \exp(-\frac{3.315}{(|\nabla L_\sigma|/k)^2}) & , |\nabla L_\sigma|_2 > 0 \end{cases}, \tag{7}$$

However, there are no analytical solutions for the PDEs involved in nonlinear diffusion filtering. The AOS algorithm uses numerical methods to approximate differential equations [24]. The diffusion equation is a semi-implicit scheme. In vector matrix representation, the discretization of Equation (5) can be expressed by Equation (8):

$$\frac{L^{i+1} - L^i}{\tau} = \sum_{i=1}^{m} A_l(L^i)L^{i+1}, \tag{8}$$

where A_l is a matrix that encodes the image transmission of each dimension. The solution of L^{i+1} can be obtained in the following way:

$$L^{i+1} = (I - \tau \sum_{i=1}^{m} A_l L^i)^{-1} L^i, \tag{9}$$

The AOS algorithm is absolutely stable for any step size. Furthermore, it also creates a discrete nonlinear diffusion scale space for any large time step. The semi-implicit method greatly improves the convergence performance by reducing mask synthesis to successive one-dimensional updates represented by tridiagonal linear equations [25].

3. Methods

This section presents our proposed stitching method completely. Aiming at the problems of inaccurate matching and obvious splicing seams in stitched images, we propose an efficacious stitching method based on color difference and an improved KAZE with a fast guided filter. Figure 1 describes the overall process of the proposed method. Allowing for the phenomenon that feature points detected by KAZE are commonly large and intensive, we introduce a fast guide filter which uses one image as a guide map to filter the other image. This method can efficiently diminish the mismatching point pairs with preservation of original visual features. During feature matching, the KAZE algorithm based on RANSAC is utilized to match the feature points which further eliminates the mismatched points. Since the color and luminance of the images may be different, there are several clear seams in the stitching result. In order to improve the splicing effect, we produce a compensation method based on color and brightness to calculate the color difference in the overlapping area. Then, the obtained local color and luminance compensation values are applied to the overall image to be spliced. In addition, our proposed method can realize the stitching of multiple images. In order to avoid excessive variation, we use the middle image as the reference image to warp the images.

3.1. Fast Guided Filtering

The first image is imported as a guide map to repair the texture information in the other image. The purpose of introducing fast guided filtering is to decline the mismatch rate of the KAZE algorithm. Here, we select a window size of 8 pixels and set the regularization parameter that controls the smoothness level to 0.01. The sampling ratio is set to four pixels in the fast guided filter. Figure 2 shows the results of fast guided filtering.

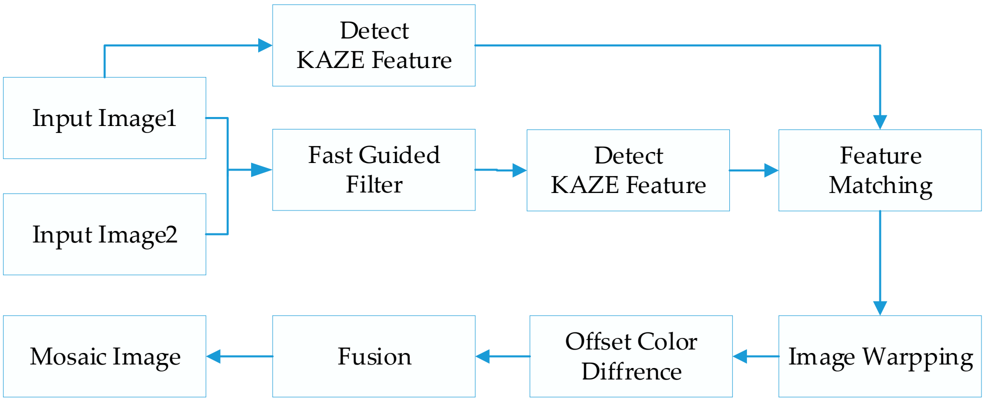


Figure 1. The proposed framework.

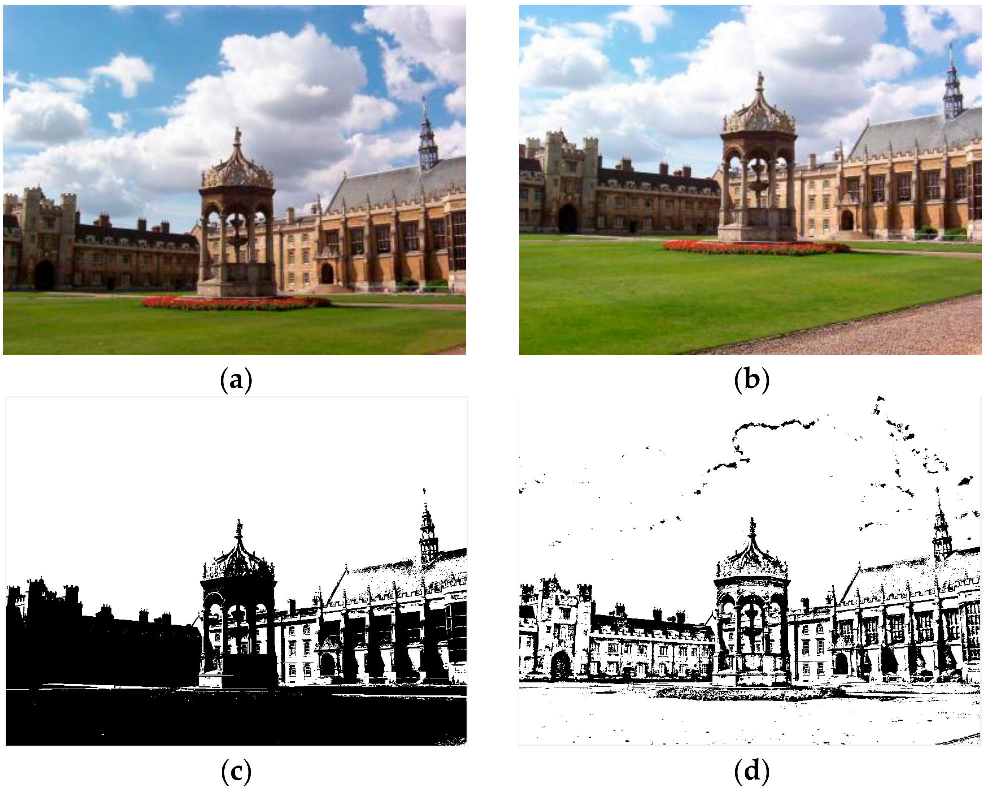


Figure 2. (**a**) is the original image; (**b**) is the guide image; (**c**) is the binary image without filtering; (**d**) is the filtered binary image.

Figure 2a,b contains input images, where (a) is the original image and (b) is the guide image. The fast guided filter filters redundant information by analyzing common textures and spatial features in both the original image and guide image. Additionally, then the fast guided filter can restore the boundaries while smoothing the original image by the guide map. This method greatly reduces the number of feature matching pairs, which results in a lower mismatch rate. After guided filtering, a lot of invalid information in the image can be filtered out, which widely improves the iterative efficiency during matching. At the same time, it can also reduce the mismatch rate to a certain extent.

3.2. KAZE Algorithm Based on RANSAC

Both the SIFT algorithm and the SURF algorithm detect feature points in a linear scale space, which can notably cause blurred boundaries and loss of detail. The KAZE algorithm detects feature points in the nonlinear scale space by constructing the nonlinear scale space, which retains more image detail. In the KAZE algorithm, AOS technology and variable transfer diffusion are used to establish a nonlinear scale space for the input images. Then, the 2D features of interest that exhibit the maximum value of scale-normalized determinant detection of Hessian response are detected by the nonlinear scale space. Finally, the scale and rotation invariance scriptors considering the first-order image derivatives are obtained through calculating the principal direction of the key point.

We introduce the RANSAC algorithm on the basis of the traditional KAZE algorithm, which can further eliminate mismatched point pairs. The flow of the KAZE algorithm based on RANSAC is shown in Figure 3.

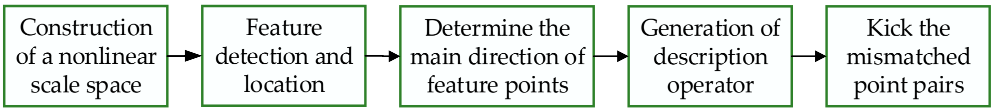

Construction of a nonlinear scale space
Feature detection and location
Determine the main direction of feature points
Generation of description operator
Kick the mismatched point pairs

Figure 3. The flow of KAZE algorithm based on RANSAC.

3.2.1. Construction of Nonlinear Scale Space

The KAZE algorithm constructs a nonlinear scale space through nonlinear diffusion filtering and the AOS algorithm [26]. The scale of KAZE features increases logarithmically. Each level of the KAZE algorithm adopts the same resolution as the original image. According to the principle of the difference of Gaussian (DoG) pyramid model [27], each group needs $S + 2$ layers of images to detect the extreme points of S scales. Then, the parameter of scale space, θ, can be described as follows:

$$\theta(o,s) = \theta_0 2^{o+\frac{s}{S}}, o \in [0,1,\ldots,O-1], s \in [0,1,\ldots,S+2], \tag{10}$$

where θ_0 is the baseline scale, o is the index of the group octave and S represents the index of the intra-group layer. The scale parameters of key points are calculated according to the group of key points and the number of layers in the group in combination with Equation (10).

The scale parameters of each layer in the group for constructing the Gaussian pyramid are calculated according to the following formulas:

$$\theta(s) = \sqrt{(r^s\theta_0)^2 - (r^{s-1}\theta_0)^2}, \tag{11}$$

$$r = 2^{\frac{1}{S}}, \tag{12}$$

From the above formulas, the scale parameters of the same layer in different groups are the same. The scale calculation formula of a layer of images in the group is shown in Equation (13).

$$\theta(s) = \theta_0 2^{\frac{s}{S}}, s \in [0,1,\ldots,S+2], \tag{13}$$

To ensure the continuity of the scale space, the first image of a group is obtained by sampling the penultimate layer of the previous group. We assume that the initial scale of the first group is θ. Then, the scale parameters of each layer in the first group are θ, $r\theta$, $r^2\theta$ and so on. The scale of the penultimate layer can be defined as follows:

$$r^S\theta = 2\theta, \tag{14}$$

Since the nonlinear diffusion filter model is based on time, it is necessary to convert the scale parameters into evolution time. We suppose that the standard deviation used in Gaussian scale space is σ. Then, the convolution of the Gaussian is equivalent to filtering the image with a duration of $t = \sigma^2/2$. According to a set of evolution time, all images in the nonlinear scale space can be obtained by the AOS algorithm as follows:

$$L^{i+1} = [I - (t_{i+1} - t_i) \sum_{i=1}^{m} A_i(L^i)]^{-1} L^i, \tag{15}$$

3.2.2. Detection and Location of Feature Points

The feature points are obtained by seeking the local maximum points of the Hessian determinant normalized by different scales. For multi-scale feature detection, we normalize the set of differential operators for scale:

$$L_{Hessian} = \theta^2 (L_{xx} L_{yy} - L_{xy}^2), \tag{16}$$

where L_{xx} and L_{yy} are the second-order horizontal derivative and vertical derivative, respectively, and L_{xy} expresses the second-order cross derivative. For the filtered image set in a nonlinear scale space, we analyze the response of the detector at different scale levels θ_i. The maximum values of scale and spatial position are searched in all filtered images except $i = 0$ and $i = N$. Then, we check the response in a 3 × 3 pixels window to quickly find the maximum value.

As shown in Figure 4, each feature point is compared with 26 pixel points from the same layer and two adjacent layers. All pixels are traversed in this way until the maximum point is found. The derivative is calculated by the Taylor expression.

$$L(x) = L + (\frac{\partial L}{\partial x})^T x + \frac{x^T}{2} (\frac{\partial^2 L}{\partial^2 x}) x, \tag{17}$$

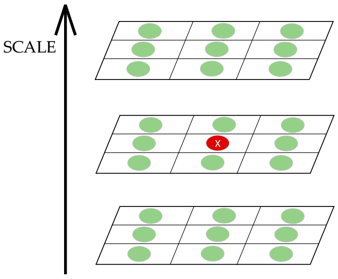


Figure 4. Detection and location of each feature point.

We take the derivative of the above equation and make it equal to zero. Then, the solution can be obtained by the following.

$$\hat{x} = (\frac{\partial^2 L}{\partial x^2})^{-1} \frac{\partial L}{\partial x}, \tag{18}$$

$$L(\hat{X}) = L + \frac{1}{2} (\frac{\partial L}{\partial x})^T \hat{X}, \tag{19}$$

One of the feature points that meets $L(\hat{X}) \geq T$ is selected as the key point.

3.2.3. Determination of Main Direction

Assuming that the scale parameter of the characteristic point is θ_i, the search radius is set to $6\theta_i$. We make six 60 degree sector regions in this radius and count the sum of Haar wavelet features in each sector region. The direction with the largest sum of wavelet features is the main direction.

3.2.4. Generation of Feature Descriptors

For feature points whose scale parameter is θ_i, we take a window of $24\theta_i \times 24\theta_i$ in the gradient image centered on the feature point. Then, the window is divided into 4×4 sub-regions with a size of $9\theta_i \times 9\theta_i$. Adjacent sub-areas have an overlapping zone with a width of $2\theta_i$. In the process of calculating the sub-region description vector, a Gaussian kernel with a weight of $\theta_1 = 2.5\theta_i$ is used for each sub-region to obtain the description vector as Equation (20):

$$d_v = (\sum L_x, \sum L_y, \sum |L_x|, \sum |L_y|), \tag{20}$$

After that, the vector d_v of each sub-region is weighted by another 4×4 Gaussian window where $\theta_2 = 1.5\theta_i$. Finally, a 64-dimensional description vector is attained through normalization processing.

3.2.5. Eliminate Mismatched Point Pairs

In the case of information loss caused by image blur, noise interference and compression reconstruction, the robustness of KAZE feature point detection is significantly better than other features. In addition, the nonlinear scale space does not cause boundary blurring and detail loss compared to the linear scale space. However, the matching of KAZE features is sensitive to the setting of parameters, which can easily prompt mismatching. Aiming at the problem of excessive concentration and mismatch of feature points, the RANSAC method is introduced to effectively eliminate mismatched points. Firstly, KAZE feature matching is performed on the previous input image and the filtered image. Additionally, then RANSAC algorithm is applied to eliminate the mismatched point pairs.

The RANSAC algorithm is used to find an optimal homography matrix. The standardized matrix satisfies Equation (21).

$$s \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \tag{21}$$

To calculate the homography matrix, we randomly extract four samples from the matching dataset which are not collinear. Then, the model is utilized for all datasets to calculate the number of points and the projection error. If the model is optimal, the loss function is the smallest. The loss function can be calculated by Equation (22).

$$\sum_{i=1}^{n} x_i' \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} + (y_i' \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}})^2, \tag{22}$$

The process of the RANSAC algorithm is shown in Algorithm 1.

Algorithm 1: The process of RANSAC

```
Input: The feature point pair S of the image to be spliced and the maximum iteration number k.
Output: Remove the feature point pair D of mismatched point pairs.
1: S->D.
2: Repeat:
3: Randomly extract four sample data from all feature point pairs detected, which cannot be
collinear.
4: Use Equations (21) and (22) to calculate the transformation matrix H, and record it as model M.
5: Calculate the projection error between all data in the dataset and model M. If the error is less
than the threshold, add the interior point set I.
6: If the number of elements of the current interior point set I is greater than the optimal interior
point set I_ Best, update I_ Best = I, update the number of iterations at the same time
$k = \frac{\log(1-p)}{\log(1-\omega^m)}$.
7: Until iterations are greater than the iteration number k.
```

The iteration number k is constantly updated rather than fixed when it is not greater than the maximum iteration number. The formula for calculation of the iteration number is as follows:

$$k = \frac{\log(1-p)}{\log(1-\omega^m)}, \tag{23}$$

where p expresses the confidence level, ω is the proportion of the interior point and m is the minimum number of samples required to calculate the model.

The RANSAC algorithm performs precise matching of feature points through the above iterative process, which effectively eliminates mismatched point pairs. The inner point pairs obtained by the method are the most advantageous pairs. The matching accuracy is the foundation for accurately evaluating warped equations. If there are many mismatched pairs, it is easy for problems such as ghosting and distortion to occur in the splicing result.

3.3. Color Difference Compensation

The luminance and color of the image to be stitched may be uneven, which can easily result in obvious stitching seams in the stitching result. To improve this problem, we put forward a color difference compensation method. The process is shown in Figure 5.

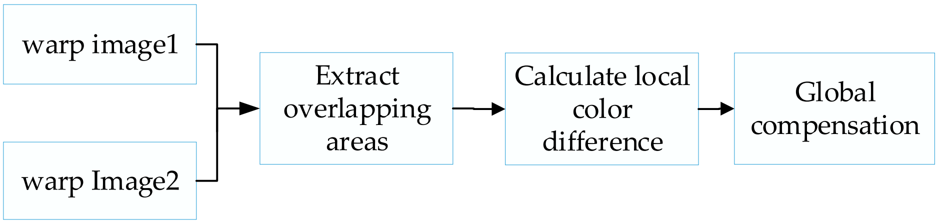


Figure 5. The flow of the color difference compensation.

We calculate the color difference and luminance difference of the overlapping area of the two distorted images. Here, we chose the LAB color model. The average color difference is calculated as follows:

$$L_{avg} = \frac{1}{n-1}\sum_{i=1}^{n}(L_{i1} - L_{i2}), \tag{24}$$

$$A_{avg} = \frac{1}{n-1}\sum_{i=1}^{n}(A_{i1} - A_{i2}), \tag{25}$$

$$B_{avg} = \frac{1}{n-1}\sum_{i=1}^{n}(B_{i1} - B_{i2}), \tag{26}$$

where L_{avg} represents the average brightness difference, L_{i1} is the brightness of the overlapping area of the first image and L_{i2} is the brightness of the overlapping area of the second image. Similarly, A_{avg} and B_{avg} are the average color difference, A_{i1} and B_{i1} are the colors of the overlapping area of the first image and A_{i2} and B_{i2} are the colors of the overlapping area of the second image. We chose the image with the higher brightness as the standard to compensate for the global brightness and color of another image. Finally, the compensated distorted image is fused to obtain the final mosaic image. Figure 6 shows the original images and Figure 7 shows the comparison of color compensated and uncompensated splicing results.

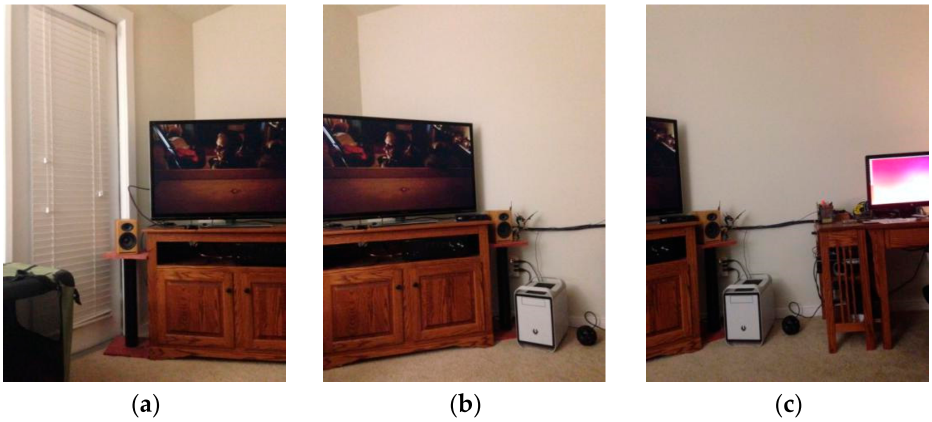

(a) (b) (c)

Figure 6. Original images. (**a**–**c**) are the original images.

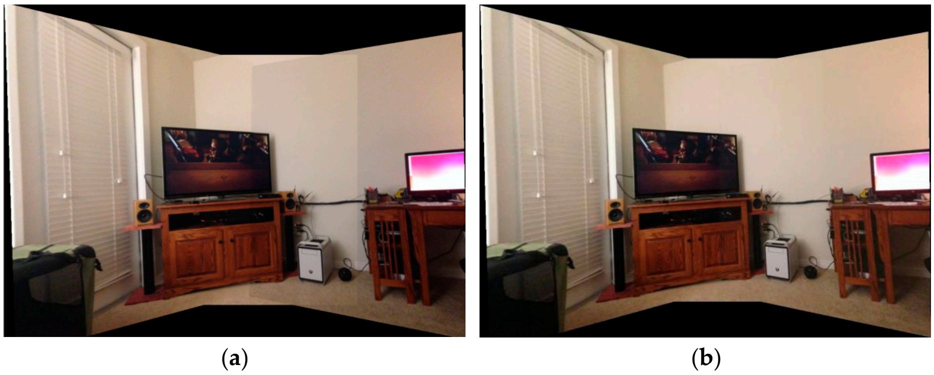

(a) (b)

Figure 7. (**a**) Mosaic image without color compensation. (**b**) Mosaic image with color compensation.

It can be seen from the figure that the stitching effect after color compensation is smoother. Not only does the stitching result have no obvious seams, but also has an improved uniformity.

4. Experiments

This section introduces some experiments to evaluate our proposed splicing method. The images used in the experiment are mainly from the ground truth database of the University of Washington [28] and the USI-SIPI image database of the University of Southern California [29]. The ground truth database of the University of Washington addresses the need for experimental data to quantitatively evaluate emerging algorithms. The high-quality and high-resolution color images in the database represent valuable extended duration digitized footage to those interested in driving scenarios or ego-motion. The images in the USI-SIPI image database of the University of Southern California have been provided for research purposes. The USI-SIPI image database contains multiple types of image sets, including texture images, aerial images, miscellaneous images and sequence images. Most of the material was scanned many years ago in the research group from a variety of sources.

At the same time, all test methods are run on a 2.60 GHz CPU with 16 GB RAM under the same experimental settings in this paper. In order to better evaluate the algorithm, we compared subjective images and quantitative evaluation. We compare the results with several popular splicing methods.

4.1. Intuitive Effect

We took a group of images in the USI-SIPI image database of the University of Southern California as an example to compare our method with SURF, BRISK, Harris, MinEigen, MSER and ORB matching algorithms. The splicing effect is shown in Table 1.

Table 1. Splicing effect.

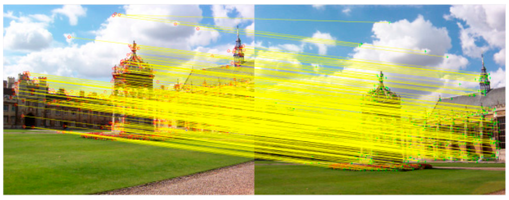
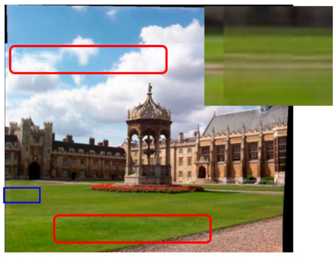
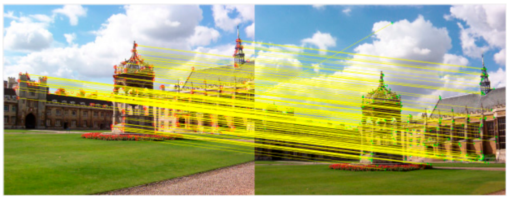
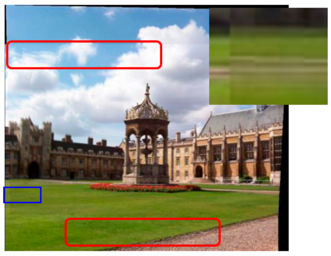
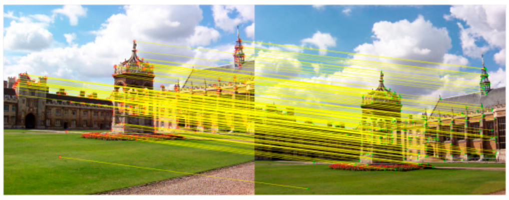
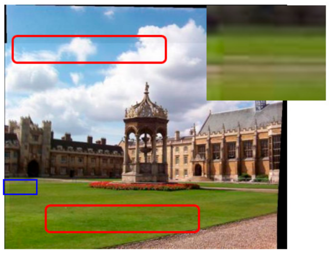
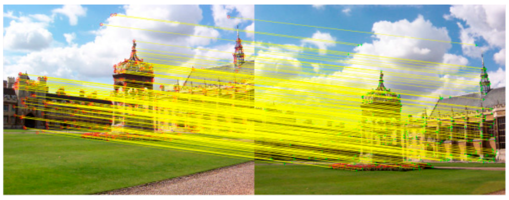
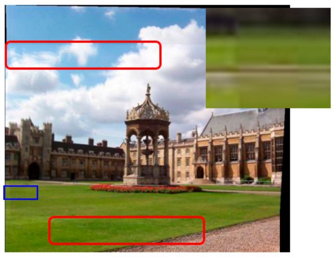
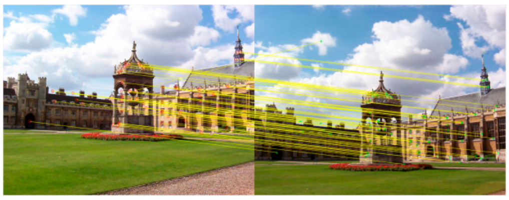
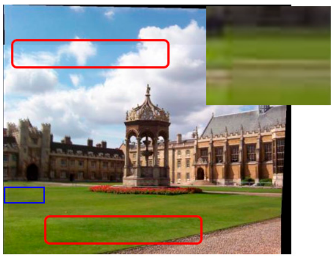
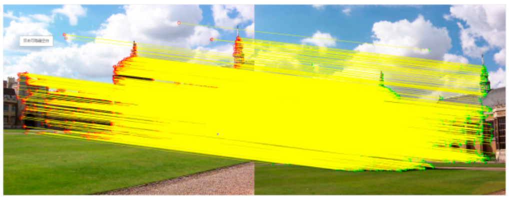
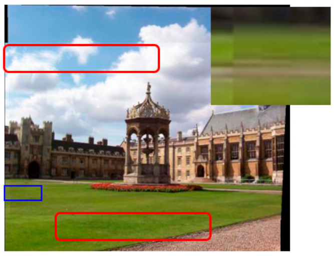
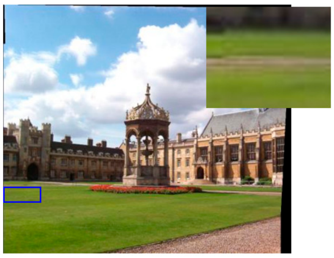
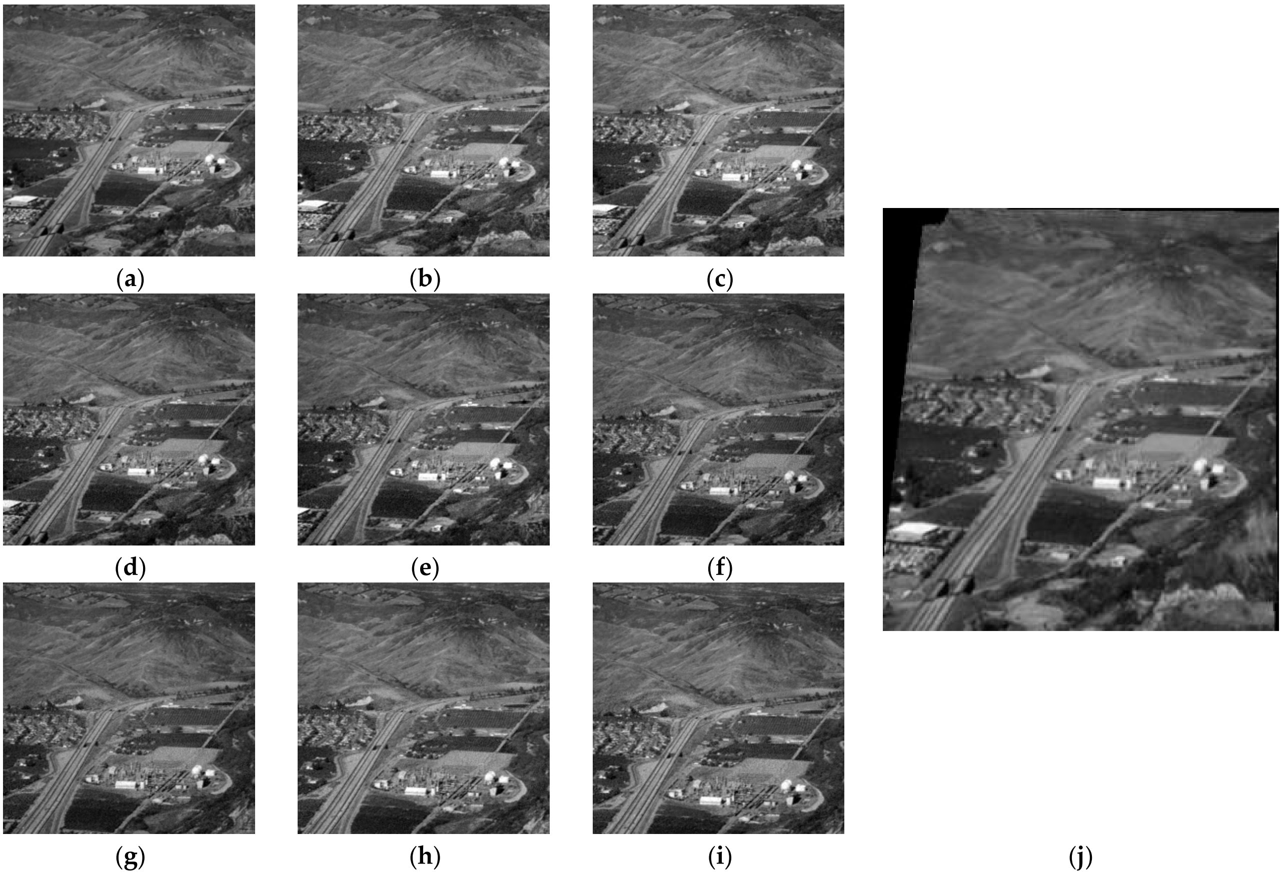
Methods	Match Figure	Stitched Image
SURF [9]		
BRISK [11]		
Harris [14]		
MinEigen		
MSER [13]		
ORB [12]		

Table 1. *Cont.*

Methods	Match Figure	Stitched Image
Our method		

The upper right corner of the stitched image in the table shows a magnification of the area marked in blue. It is not difficult to see from the table that the proposed algorithm is not only better than other algorithms regarding detail, but also more uniform in the color and brightness of the mosaic image. As shown in the red marked area in the table, the stitching method we proposed has no obvious stitching seams, while other algorithms have several obvious stitching seams.

In addition, our method can be applied to multiple image stitching. We selected multiple UAV images from the image library, and the splicing effect is shown in Figure 8.



(**a**) (**b**) (**c**) (**d**) (**e**) (**f**) (**g**) (**h**) (**i**) (**j**)

Figure 8. UAV images panoramic mosaic: (**a**–**i**) are the original images; (**j**) is the stitched image.

4.2. Quantitative Evaluation

In order to effectively evaluate the effectiveness of the algorithm, we quantitatively evaluated the above algorithms from four aspects: the number of matching pairs, the correct matching rate, the root mean square error (RMSE) and the mean absolute error (MAE) [30–32].

Assuming that the total number of matches is N and the correct matching logarithm is A, the correct matching rate can be defined as Equation (27).

$$\text{Accurancy} \;=\; \frac{\text{A}}{\text{N}} \times 100\%, \tag{27}$$

We randomly selected eight groups (Table 2) of data from the two databases for feature matching. The table shows the matching logarithms and the matching accuracy results of several comparison algorithms.

Table 2. The results of a comparison of different algorithms.

Data	Estimate	SURF	BRASK	MSER	MinEigen	Harris	ORB	Proposed
Data1	Matching pairs Accuracy	69 89.85%	14 **100%**	20 90.00%	32 96.87%	19 **100%**	101 99.01%	267 98.50%
Data2	Matching pairs Accuracy	123 99.18%	71 98.59%	19 94.73%8	68 **100%**	36 **100%**	321 99.38%	243 99.59%
Data3	Matching pairs Accuracy	205 98.54%	109 99.08%	51 96.08%	125 100%	87 97.70%	771 99.48%	606 **100%**
Data4	Matching pairs Accuracy	474 99.36%	231 98.70%	135 95.56%	232 98.28%	157 98.73%	2348 99.32%	1280 **99.38%**
Data5	Matching pairs Accuracy	61 90.16%	38 100%	7 71.43%	39 100%	22 95.45%	167 98.80%	148 **100%**
Data6	Matching pairs Accuracy	39 92.31%	15 **100%**	10 **100%**	29 96.55%	19 **100%**	133 99.93%	154 99.35%
Data7	Matching pairs Accuracy	564 97.69%	266 99.25%	79 96.20%	398 99.50%	233 98.71%	3112 99.94%	1242 **100%**
Data8	Matching pairs Accuracy	154 94.81%	44 88.64%	25 76.00%	104 91.35%	54 92.59%	440 97.95%	212 **98.11%**

It can be seen from the above table that the extent of feature matching based on ORB is the largest. Although the overall matching of the ORB is relatively stable, the matching rate is lower than the proposed method in this paper. Other comparison algorithms detect fewer feature points. Even though other algorithms have high matching rates for individual data, their performance in terms of the number of feature points, matching accuracy and stability is poor. The improved KAZE algorithm based on a fast guided filter is second only to the ORB algorithm in the number of feature matching pairs. Our proposed method further filters out the mismatched point pairs, thereby improving the matching efficiency and matching stability.

In order to judge the stability of the matching algorithm more intuitively, we compared the accuracy of several algorithms in a line chart (Figure 9).

The red color in the figure shows the matching accuracy of our proposed algorithm. It is not difficult to see from the figure that the algorithm we proposed is not only more accurate, but also more stable.

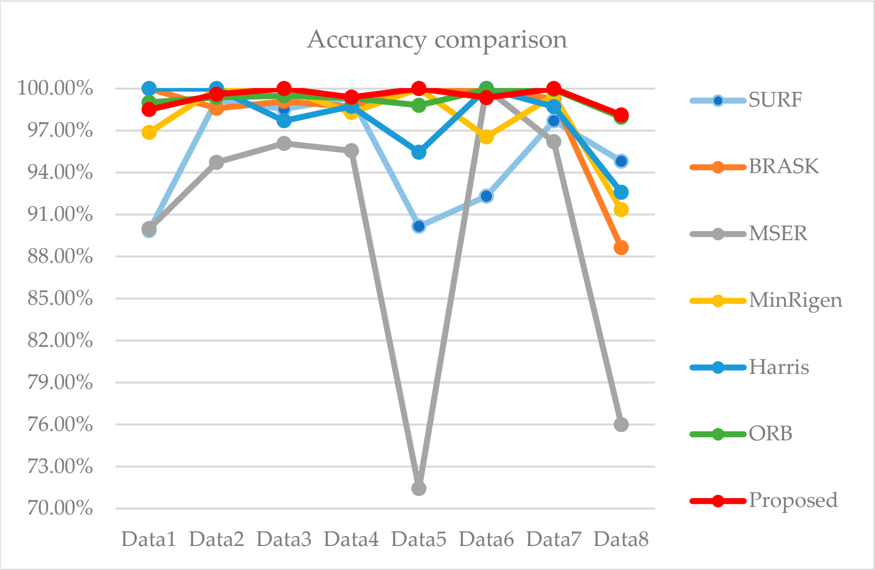


Figure 9. Comparison of accuracy.

For two images, I_1 and I_2, to be spliced, given N matching pairs (pi1 and pi2), where i = 1, 2, N, the RMSE and the MAE are defined as follows:

$$\mathrm{RMSE}(\mathrm{I},\mathrm{I}') = \sqrt{\frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}(\mathrm{f}(\mathrm{p_i}) - \mathrm{p}_i')^2}, \tag{28}$$

$$\mathrm{MAE}(\mathrm{I},\mathrm{I}') = \frac{1}{\mathrm{N}}\sum_{i=1}^{\mathrm{N}}|\mathrm{f}(\mathrm{p_i}) - \mathrm{p}_i'|, \tag{29}$$

Additionally, taking the selected eight groups (Table 3) of data as an example for splicing, the table shows the comparison results of the RMSE and the MAE of several different algorithms.

Table 3. The comparison results of the RMSE and the MAE of different algorithms.

Data	Estimate	SURF	BRASK	MSER	MinEigen	Harris	ORB	Proposed
Data1	RMSE MAE	**0.1037** **0.0759**	0.1354 0.1157	0.1301 0.1051	0.1321 0.1060	0.1309 0.1096	0.1261 0.0935	0.1237 0.0935
Data2	RMSE MAE	0.0917 0.0671	0.1367 0.0933	0.0973 0.0819	0.1173 0.0916	0.1185 0.0846	0.1099 0.0833	**0.0797** **0.0649**
Data3	RMSE MAE	0.1285 **0.0992**	0.1439 0.1155	0.1386 0.1243	0.1738 0.1370	0.1727 0.1387	0.1831 0.1505	**0.1283** 0.0999
Data4	RMSE MAE	0.0920 0.0592	0.0946 0.0655	0.0966 0.0597	0.1616 0.1124	0.1734 0.1194	0.1417 0.0905	**0.0903** **0.0545**
Data5	RMSE MAE	0.0776 0.0588	0.0895 0.0650	0.2573 0.1305	0.0867 0.0708	0.8250 0.0717	0.1169 0.0769	**0.0760** **0.0586**
Data6	RMSE MAE	0.0539 0.0398	0.0713 0.0510	0.0562 0.0467	0.0599 0.0429	0.0482 0.0394	0.0691 0.0458	**0.0417** **0.0302**
Data7	RMSE MAE	0.1074 0.0847	0.1120 0.0888	0.1386 0.1084	0.1392 0.1040	0.1222 0.0947	0.1494 0.1075	**0.1038** **0.0842**
Data8	RMSE MAE	0.0705 0.0494	0.0719 0.0496	0.0400 0.0345	0.1139 0.0777	0.1047 0.0812	0.0994 0.0666	**0.0701** **0.0447**

The table shows that the RMSE and the MAE values of our proposed method are lower. The overall stitching effect is better than other stitching algorithms.

In addition, we also compared the processing time of the proposed algorithm with other algorithms. The result shows that the processing time of the proposed algorithm is similar to that of other algorithms. The comparison results of processing time are shown in Table 4.

Table 4. The comparison results of processing time of different algorithms.

Data	SURF	BRASK	MSER	MinEigen	Harris	ORB	Proposed
Data1	**1.226 s**	1.681 s	1.456 s	1.772 s	1.354 s	1.879 s	1.798 s
Data2	2.239 s	3.001 s	2.489 s	2.096 s	**2.204 s**	5.370 s	3.232 s
Data3	1.168 s	**1.444 s**	1.549 s	2.238 s	1.179 s	2.523 s	2.389 s
Data4	**1.219 s**	1.341 s	1.318 s	1.996 s	1.487 s	2.653 s	1.979 s
Data5	**1.149 s**	1.362 s	1.251 s	1.359 s	1.160 s	2.568 s	1.356 s
Data6	1.715 s	2.406 s	1.998 s	1.337 s	**1.181 s**	1.661 s	1.509 s
Data7	**1.089 s**	1.964 s	1.981 s	1.549 s	1.347 s	2.017 s	1.973 s
Data8	**2.147 s**	3.010 s	2.984 s	2.471 s	2.309 s	4.175 s	2.992 s

It is not difficult to draw the conclusion from Table 4 that the processing time of the SURF algorithm is relatively low and the ORB algorithm is the most time consuming among the studied algorithms. The proposed method has similar processing times compared to other methods.

5. Conclusions

In this paper, we propose an improved image stitching algorithm based on color difference and KAZE with a fast guided filter, which solves the problems of a high mismatch rate in feature matching and obvious seams in the stitched image. In this paper, a fast guide filter is introduced to reduce the mismatch rate before matching. The KAZE algorithm based on RANSAC is used for feature matching of the image to be spliced, and the matrix transformation of the image to be spliced is performed. Then, the color difference and brightness difference of the overlapping area of the image to be spliced is calculated, and overall adjustments of the image to be spliced are made so as to improve the nonuniformity of the spliced image. Finally, the converted image is fused to obtain the stitched image.

Our proposed method was evaluated via the resulting visual image and quantitative value of the images of the ground truth database and USI-SIPI image database and compared with other popular algorithms. On the one hand, the stitching method we proposed can achieve smoother and more detailed stitched images. On the other hand, the algorithm proposed in this paper is superior to other algorithms in terms of matching accuracy, the RMSE and the MAE.

The proposed stitching method in this paper can be used in many fields, such as UAV panoramic stitching and virtual reality. However, the limitations of our proposed method are as follows: firstly, the method proposed in this paper is prone to ghosting when splicing moving objects in close proximity. Secondly, the proposed algorithm has a similar processing time compared to other algorithms, but cannot achieve the effect of real-time stitching. In the future, we will further improve the efficiency of image stitching.

Author Contributions: Conceptualization, C.Z.; methodology, C.Z., D.W. and H.S.; formal analysis, C.Z. and D.W.; investigation, C.Z. and D.W.; resources, H.S., C.Z. and D.W.; writing—original draft preparation, C.Z.; writing—review and editing, C.Z.; supervision, D.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Youth Foundation of Jilin Province 62205332.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gómez-Reyes, J.K.; Benítez-Rangel, J.P.; Morales-Hernández, L.A. Image Mosaicing Applied on UAVs Survey. *Appl. Sci.* **2022**, *12*, 2729. [CrossRef]
2. Jiang, X.; Ma, J.; Fan, A. Robust feature matching for remote sensing image registration via linear adaptive filtering. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 1577–1591. [CrossRef]
3. Liao, T.; Li, N. Single-perspective warps in natural image stitching. *IEEE Trans. Image Process.* **2019**, *29*, 724–735. [CrossRef] [PubMed]
4. Nie, L.; Lin, C.; Liao, K. Learning edge-preserved image stitching from multi-scale deep homography. *Neurocomputing* **2022**, *491*, 533–543. [CrossRef]
5. Lin, M.; Liu, T.; Li, Y. Image stitching by disparity-guided multi-plane alignment. *Signal Process.* **2022**, *197*, 108534. [CrossRef]
6. Zhang, L.; Huang, H. Image Stitching with Manifold Optimization. *IEEE Trans. Multimed.* **2022**, *10*, 1109. [CrossRef]
7. Zhang, Y.; Lai, Y.K.; Zhang, F.L. Content-preserving image stitching with piecewise rectangular boundary constraints. *IEEE Trans. Vis. Comput. Graph.* **2020**, *27*, 3198–3212. [CrossRef]
8. Li, Y.; Yang, C.; Zhang, L. A novel SURF based on a unified model of appearance and motion-variation. *IEEE Access* **2018**, *6*, 31065–31076. [CrossRef]
9. Liu, J.; Bu, F.L. Improved RANSAC features image-matching method based on SURF. *J. Eng.* **2019**, *2019*, 9118–9122. [CrossRef]
10. Guan, H.; Smith, W.A.P. BRISKS: Binary features for spherical images on a geodesic grid. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4516–4524.
11. Liu, Y.; Zhang, H.; Guo, H. A fast-brisk feature detector with depth information. *Sensors* **2018**, *18*, 3908. [CrossRef]
12. Zhang, M.Y.; Wang, Y.Y.; Hou, X.H. Image Stitching Algorithm Based on ORB and Improved RANSAC. *Comput. Sci.* **2019**, *46*, 294–298. [CrossRef]
13. Ordóñez, Á.; Heras, D.B.; Argüello, F. Exploring the MSER-based hyperspectral remote sensing image registration. In Proceedings of the Image and Signal Processing for Remote Sensing XXVI, SPIE, Online, 21–25 September 2020; Volume 11533, pp. 89–96.
14. Elgamal, Z.M.; Yasin, N.B.M.; Tubishat, M. An improved harris hawks optimization algorithm with simulated annealing for feature selection in the medical field. *IEEE Access* **2020**, *8*, 186638–186652. [CrossRef]
15. Debnath, A.; Das, S. Identification of Image Forgery based on various Corner Detection methods. In Proceedings of the 2022 IEEE World Conference on Applied Intelligence and Computing (AIC), Sonbhadra, India, 17–19 June 2022; pp. 314–319.
16. 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]
17. Alcantarilla, P.F.; Bartoli, A.; Davison, A.J. KAZE features. In Proceedings of the Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, 7–13 October 2012; pp. 214–227.
18. Pourfard, M.; Hosseinian, T.; Saeidi, R. KAZE-SAR: SAR image registration using KAZE detector and modified SURF descriptor for tackling speckle noise. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5207612. [CrossRef]
19. Li, Z.; Zheng, J.; Zhu, Z.; Yao, W.; Wu, S. Weighted guided image filtering. *IEEE Trans. Image Process.* **2014**, *24*, 120–129.
20. He, K.; Sun, J. Fast guided filter. *arXiv* **2015**, arXiv:1505.00996.
21. Fan, J.; Wu, Y.; Li, M.; Liang, W.; Cao, Y. SAR and optical image registration using nonlinear diffusion and phase congruency structural descriptor. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 5368–5379. [CrossRef]
22. Weickert, J.; Romeny, B.T.H.; Viergever, M.A. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Trans. Image Process.* **1998**, *7*, 398–410. [CrossRef]
23. Li, Y.; Ding, Y.; Li, T. Nonlinear diffusion filtering for peak-preserving smoothing of a spectrum signal. *Chemom. Intell. Lab. Syst.* **2016**, *156*, 157–165. [CrossRef]
24. Jeon, M.; Alexander, M.; Pedrycz, W. Unsupervised hierarchical image segmentation with level set and additive operator splitting. *Pattern Recognit. Lett.* **2005**, *26*, 1461–1469. [CrossRef]
25. Shen, Y.; Peng, F.; Zhang, Z. Efficient optical proximity correction based on semi-implicit additive operator splitting. *Opt. Express* **2019**, *27*, 1520–1528. [CrossRef] [PubMed]
26. Jong, T.K.; Bong, D.B.L. An effective feature detection approach for image stitching of near-uniform scenes. *Signal Process. Image Commun.* **2023**, *110*, 116872. [CrossRef]
27. Setumin, S.; Suandi, S.A. Difference of Gaussian oriented gradient histogram for face sketch to photo matching. *IEEE Access* **2018**, *6*, 39344–39352. [CrossRef]
28. Brostow, G.J.; Fauqueur, J.; Cipolla, R. Semantic object classes in video: A high-definition ground truth database. *Pattern Recognit. Lett.* **2009**, *30*, 88–97. [CrossRef]

29. Mahto, D.K.; Anand, A.; Singh, A.K. Hybrid optimisation-based robust watermarking using denoising convolutional neural network. *Soft Comput.* **2022**, *26*, 8105–8116. [CrossRef]
30. Wang, Y.; Yuan, Y.; Lei, Z. Fast SIFT feature matching algorithm based on geometric transformation. *IEEE Access* **2020**, *8*, 88133–88140. [CrossRef]
31. Khajehzadeh, M.; Taha, M.R.; Keawsawasvong, S. An effective artificial intelligence approach for slope stability evaluation. *IEEE Access* **2022**, *10*, 5660–5671. [CrossRef]
32. Qi, J.; Du, J.; Siniscalchi, S.M. On mean absolute error for deep neural network based vector-to-vector regression. *IEEE Signal Process. Lett.* **2020**, *27*, 1485–1489. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.