

Article

Utilization of Merge-Sorting Method to Improve Stitching Efficiency in Multi-Scene Image Stitching

Wei Liu ^{1,†}, Kunhao Zhang ^{2,†}, Yan Zhang ¹, Jian He ² and Bo Sun ^{2,*} ¹ Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China² Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201204, China

* Correspondence: sunb@sari.ac.cn

† These authors contributed equally to this work.

Abstract: Aiming at the problem of single stitching scenes (ordered or disordered images), which are time-consuming and involve complex computation during the image stitching, a multi-scene image stitching method based on merge sorting is proposed called merge-sorting image stitching (MSIS). We established an image stitching model based on merge sorting, directly stitching the matched images without traversing all the images and sorting them. Compared with the traditional method, this method significantly reduces the computational time taken matching irrelevant image pairs and improves the efficiency of image registration and stitching. The experimental results show that, compared with the binary tree-based model, our method reduces the time required to complete image stitching by about 10% and the distortion of the stitched images.

Keywords: image stitching; unordered images; multiple scenes; merge sorting; A-KAZE

Citation: Liu, W.; Zhang, K.; Zhang, Y.; He, J.; Sun, B. Utilization of Merge-Sorting Method to Improve Stitching Efficiency in Multi-Scene Image Stitching. *Appl. Sci.* **2023**, *13*, 2791. <https://doi.org/10.3390/app13052791>

Academic Editor: Seokwon Yeom

Received: 28 December 2022

Revised: 26 January 2023

Accepted: 4 February 2023

Published: 21 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, computer vision technology has developed rapidly, and image technology has been applied in various industries. Image technology is widely used for VR and 360 degree panoramic images. Image stitching technology is particularly important for panoramic images, and a lot of research has been carried out on image stitching algorithms. Image stitching algorithms stitch small parts of scenes together to obtain a picture of the whole scene. They are widely used in various fields, including for 3D reconstruction of objects [1], medical image analysis in medicine [2], remote sensing image analysis in geography [3], and 360 degree panoramic views outside cars in autonomous driving [4]. At present, the technology mainly focuses on the stitching of ordered images, and there is very little research on unordered image stitching. After consulting a large number of studies, it was found that unordered image stitching is similar to ordered image stitching, which involves stitching multiple images with overlapping areas into a panoramic image with a wide viewing angle and high resolution [5,6]. This technique is suitable not only for stitching ordered images but also for unordered images.

The main steps employed by image stitching technology can be divided into image preprocessing, image matching and registration, and image fusion, with the latter two steps being the core steps [7]. Image preprocessing involves inputting the images that need to be stitched into the model to prepare for subsequent image matching and fusion. Image matching is the process of determining the overlapping area between two or more images through feature point matching. The accuracy and speed of feature point matching directly affect the results and efficiency of image stitching. Image registration is the process of aligning two or more images after image matching into a single wide-field image. During the registration process, wrongly matching points are removed with the matching method, and the correctly matching points are retained to ensure that the images are stitched together nicely. Image fusion is the process of fusing the registered images into a wide-field image using an image fusion method, which directly determines the final stitching result.

If the image fusion method used in this process is not suitable, the final stitching result will not achieve the desired effect, resulting in the failure of image stitching.

The more commonly used feature point detection methods include scale-invariant feature transform (SIFT) [8], accelerated robust function (SURF) [9], oriented FAST and rotated BRIEF (ORB) [10], and A-KAZE [11]. After the feature points of the images are detected, it is necessary to match the detected feature points and then remove the incorrect feature points. Methods for image feature point matching include the brute force (BF), fast library of approximate nearest neighbors (FLANN), K-nearest neighbor (KNN) matcher, and filter random sample consensus (RANSAC) [12,13] methods. Image fusion methods include linear weighted fusion, Laplacian pyramid fusion [14], and optimal suture fusion [15,16].

In recent years, research on sequential image stitching has progressed rapidly, and many improved methods for image registration and fusion have been proposed. An et al. [17] used a combination of SURF optimization and improved cell-accelerated power function weighting to eliminate erroneous matching points in image matching. Shi et al. [18] further purified the feature points by using the sampling consistency algorithm to match the feature points. Zhu et al. [19] reduced the complexity of feature point descriptor data by using the ORB algorithm and principal component analysis (PCA) to process feature points. Lu Jin et al. [20] increased the number of surface matching points for objects in images by using adaptive histogram equalization and used the improved zero-mean normalized cross-correlation (ZNCC) algorithm to reduce the false matching rate for feature points and improve the matching efficiency. Huang Hua et al. [21] used the effective part of second-order minimization (ESM) to calculate geometric errors and reduce ghosting and distortion, employing interpolation between homography and similarity transformations with a general matrix manifold. Chen et al. [22] used a Gaussian pyramid combined with a Laplacian pyramid to fuse images. The abovementioned methods for image feature point detection, image feature point matching, and image fusion are mainly used in the stitching of ordered images, and there are few studies on methods for unordered image stitching. If the stitching method used with ordered images is directly applied to the stitching of unordered images, it will have a great impact on the image feature point matching and fusion, and some may even result in errors and not be able to perform image fusion. Even if image fusion can be performed, the final stitched result may be incomplete, and the information from the two images is lost. In severe cases, the result of stitching may be completely black or the image may be severely distorted, and the information from the image will not be recognizable at all.

Unordered image stitching is very different from ordered image stitching. Ordered images are pairs of adjacent images that have overlapping areas. When stitching, the adjacent two images can be directly stitched, and the stitching order goes from left to right. Unordered images are pairs of adjacent images that may or may not have overlapping areas. Two images with overlapping areas need to be detected before stitching, and the stitching order is not fixed. Thus, compared with ordered image stitching, the efficiency of unordered image stitching is lower and the process is more time-consuming.

In order to achieve the stitching of unordered images, the traditional method involves finding image pairs in an unordered image set for which feature point matching is possible and then sorting the matched images before performing image feature point matching and image fusion. Although the stitching of unordered images can be achieved, the stitching efficiency is low and the process is time-consuming. The most time-consuming stage is finding all the image pairs that can be matched. Snavely et al. [23] used an exhaustive matching method in which, for example, if there are N images, the method must match $N(N-1)/2$ image pairs. Zhong et al. [7] first estimated the overlapping areas of the images and then used the improved binary tree model to stitch the unordered images. The above two methods for stitching unordered images can both complete the stitching and produce good results, and they also improve the stitching efficiency compared to the traditional methods to a certain extent. However, both methods need to traverse all the images or

estimate all their overlapping areas before performing image fusion. In this way, the problems of the process being time-consuming and having low efficiency arise.

In order to solve the problems of low efficiency and a time-consuming process that arise with these methods, we propose a multi-scene image stitching method based on the merge-sorting method (merge-sorting image stitching (MSIS)). Our main contributions are as follows: (1) We develop an improved model based on the merge-sorting method to perform feature point matching and image stitching with unordered images in order to finally obtain an image of the entire scene. Compared with traditional methods, this method reduces the time required for unordered image matching and fusion and improves the stitching efficiency for unordered images. (2) The experimental results show that the MSIS method can not only improve the efficiency of unordered image stitching but also that of ordered image stitching compared to other unordered image stitching methods.

The other sections in this paper are organized as follows. In Section 2, related work on image registration and image fusion is briefly introduced. In Section 3, we detail the image registration method used in this study and present the registration results. In Section 4, the image stitching method based on the merge-sorting model and the image fusion method used in this study are introduced in detail. The experimental results and a comparison of this method with other methods are presented in Section 5. Finally, Section 6 provides the conclusion of this study.

2. Related Work

2.1. Image Matching Method

In recent years, the SIFT, SURF, ORB, and A-KAZE feature point detection algorithms have been widely used. Bay et al. [24] and Chen et al. [25] improved the SIFT and SURF algorithms, respectively, increasing the computational efficiency but decreasing the feature point registration accuracy. Alcantarilla et al. [26] proposed the accelerated-KAZE (A-KAZE) algorithm in 2003, which is a matching algorithm based on nonlinear diffusion filtering and the fast explicit diffusion (FED) framework; it is also an improvement of the KAZE algorithm. The A-KAZE algorithm not only increases the grayscale diffusion efficiency for smooth images but also reduces the speed at the edges, and it can preserve the details of edges well. Therefore, we used the A-KAZE algorithm for image feature point detection and matching in this study.

2.2. Image Fusion Method

After images are registered according to the method of image registration, they are fused. The linear weighted fusion method [27] can quickly produce fused images, but the overlapping areas often appear as ghosting and the sharpness of the images is reduced during fusion. Laplacian pyramid image fusion [14] not only solves the ghosting problem but also the problem of fusing a small number of images with different exposures; however, it is computationally intensive and time-consuming. The optimal seam line fusion algorithm [16] is an algorithm based on image cutting. In the overlapping area, the cutting line with the smallest energy intensity value obtained through the energy function is the optimal suture line. The optimal seam line algorithm not only has a good fusion effect but also lower computational complexity than the Laplacian pyramid algorithm. Therefore, we used the optimal seam line fusion algorithm as the image fusion algorithm in this study.

3. Image Registration

Image registration is the core step of image stitching. The result of image registration directly affects the result of the subsequent image fusion. For the same scene, images with overlapping areas generally have overlapping areas of about 30% to 60%. If there is an overlapping area between two images, the image fusion algorithm can be used after image registration to stitch the two images into a wide-field image. Other images with overlapping areas are also registered and fused in the same way, and then they are stitched in sequence. Finally, the scene can be stitched into a complete image.

The detection of feature points by the A-KAZE algorithm can be divided into three main steps. The first step is the building of the nonlinear scale space, which is followed by the detection of feature points. The last step is the generation of feature descriptors; that is, feature point matching.

When building nonlinear scale spaces, the A-KAZE algorithm uses a nonlinear diffusion filter that often employs a divergence of a certain flow function to describe the variation in luminance in different spaces. The nonlinear partial differential equation is:

$$\frac{\partial L}{\partial t} = \operatorname{div}(c(x, y, t) \cdot \nabla L) \quad (1)$$

where div and ∇ represent the divergence and gradient, respectively; L represents the brightness of the image; c is the transfer diffusion function used to adapt to the local structure of the image to ensure its local accuracy; and t in the function c is the scale parameter of the image: the larger the value of t , the greater the scale of the image. The function c is defined as:

$$c(x, y, t) = g(\nabla L_\sigma(x, y, t)) \quad (2)$$

where g is defined as:

$$g = \frac{1}{1 + \frac{|\nabla L_\sigma|^2}{\lambda^2}} \quad (3)$$

Here, ∇L_σ represents the image gradient for the Gaussian smoothing with the image brightness as the standard deviation σ , and the parameter λ is the contrast factor that controls the degree of diffusion. The size of λ determines the boundary of the image to be enhanced or weakened during the smoothing process. The larger the value of λ is, the less boundary information is preserved and the smoother the image; in contrast, the more boundary information is preserved, the clearer the image is.

After the nonlinear scale space is constructed, the detection and localization of feature points are carried out. Similar to the SIFT algorithm, A-KAZE feature point detection uses the non-maximum suppression method to calculate the Hessian matrix for each pixel in the image pyramid. The formula for calculating the determinant of the Hessian matrix is:

$$L_{Hessian} = \sigma^2 (L_{xx} L_{yy} - L_{xy}^2) \quad (4)$$

where σ is the integer value of the scale parameter σ_i of a certain layer; L_{xx} and L_{yy} represent the second-order partial derivatives of the image in the horizontal and vertical directions, respectively; and L_{xy} represents the partial derivatives of x and y , respectively. Then, the Taylor expansion is used to determine the coordinates of the feature points, and the solution formula is as follows:

$$L(x) = L + \left(\frac{\partial L}{\partial x} \right)^T x + \frac{1}{2} x^T \frac{\partial^2 L}{\partial x^2} x \quad (5)$$

where $L(x)$ represents the spatial scale function of the image brightness, and x represents the coordinates of the feature points. The formula to solve for these coordinates is:

$$\hat{x} = - \left(\frac{\partial^2 L}{\partial x^2} \right)^{-1} \frac{\partial L}{\partial x} \quad (6)$$

Feature point matching involves constructing a KD tree of the two images to be matched and then selecting the first image as a reference image for KNN matching. The public matching pairs after two matches are used as the initial matching pairs. Then, the RANSAC algorithm is used to remove the incorrectly matched points, keeping the correctly matched points, and to calculate the affine transformation matrix for the images. After several iterations, the final image feature point matching result is obtained.

We used three methods (SIFT, SURF, and A-KAZE) to detect and match the feature points of the images respectively and compared the time taken by the three methods, which is discussed in Section 5. In Figure 1, four images (a,b) and (c,d) were obtained from [28]. It shows the results for the feature point detection and matching obtained using the A-KAZE method for the pairs of images (a,b) and (c,d), respectively. The pair of images (a,b) is a close-up view with shade, and the pair of images (c,d) is a distant view with roads and pedestrians. It can be seen from the results that the method not only resulted in correctly matching the points for buildings in all images, but matching points in the shade of trees were also correctly matched in images (a,b), and matching points on the road were correctly matched in images (c,d).

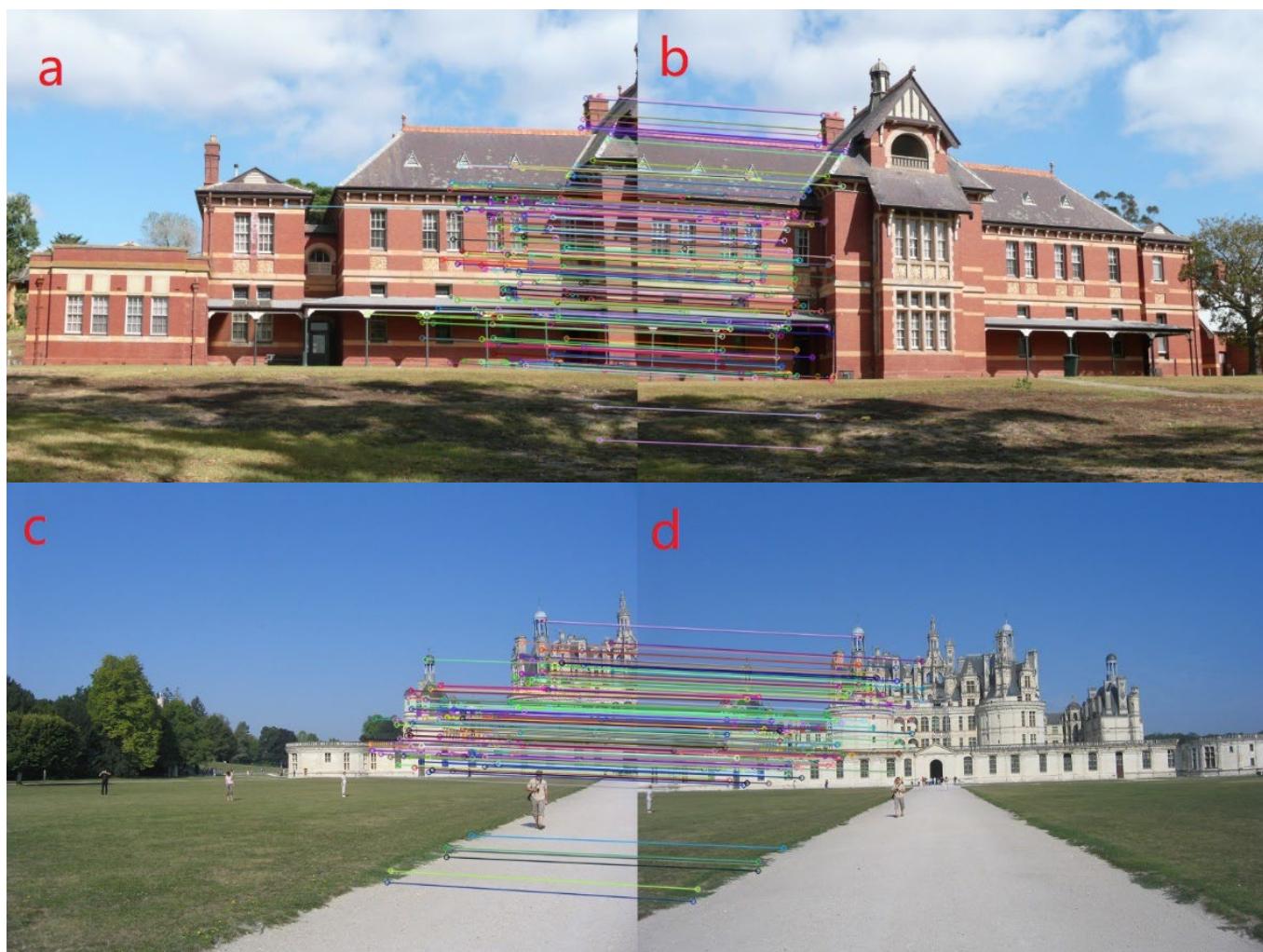


Figure 1. Pairs of subfigures, (a–d), are the results for matching image feature points using A-KAZE and KNN algorithms.

4. Multi-Scene Image Stitching Based on Merge-Sorting Method

In order to be able to stitch both ordered and unordered images and to improve the efficiency of image stitching and reduce the stitching time, this paper proposes a multi-scene image stitching model based on the merge-sorting method. A schematic diagram of the structure of the model is shown in Figure 2.

We input N images into the model and divided these N images into $N/2$ groups, then merged each of the two images into one group. This step was similar to the traditional merge-sorting model, where the images in the parent node of the upper layer are merged with the images in the two adjacent leaf nodes. At the same time, the feature points of the

merged two images were also matched. When the number of matching points reached the minimum, image stitching was performed with the upper-layer nodes, and the images that did not reach the minimum number of matching points were directly used as the upper-layer nodes. Then, we processed the images in the upper-layer nodes. In order to reduce the stitching time, we set a feature point matching condition that had to be satisfied by the image fusion: the first image was used as a reference image, and the feature points were matched with the subsequent images in turn. When the number of feature points matched in an image reached the minimum number of matching points, the two images were fused as an upper-layer-node image, instead of matching all the images once and then performing image fusion. When the two images were fused, the next image was used as the reference image, and then the feature points were matched and fused with the following images in turn. Images that underwent image fusion no longer participated in feature point matching and fusion. In this way, the feature points of the previous images and the following images were matched, and the images that met the matching condition for the feature points were fused until there were no images that met the condition for this layer of images and no images that met the condition for the images that were directly used in the parent node of the upper layer. By using this method for feature point matching and fusion with the images from each layer, we finally obtained a panoramic image at the root node. After many experiments and comparison and induction, we found that the number of feature points required to match two images with overlapping areas was generally more than 80. In order to better carry out image fusion, the minimum matching point value was set to 50 in this study.

After matching the feature points and meeting the condition for the minimum number of matching points, the next step was image fusion. In this study, the optimal seam line image fusion method was used, which involves identifying the “low energy” area in the overlapping area of two images, using the energy formula to calculate the energy value for a single pixel in this area, and then determining a path through which image fusion can be achieved with the lowest energy. This path is called the optimal seam line. Following [29], the energy formula used in this study was as follows:

$$E(i, j) = \alpha \cdot e_c(i, j) + \beta \cdot e_g(i, j) \quad (7)$$

where α and β were set to 0.83 and 0.17, following [29]. e_c is the grayscale difference map for the image and e_g is the texture structure difference map for the image, and their formulas are:

$$e_c = \left\| \frac{1}{5 \times 5} \sum_{m=-2}^2 \sum_{n=-2}^2 img_sub(i+m, j+n) - img_sub(i, j) \right\| \quad (8)$$

$$e_g = \left| \left(\frac{1}{5 \times 5} \sum_{m=-2}^2 \sum_{n=-2}^2 Grad_x(i+m, j+n) \right)^2 + \left(\frac{1}{5 \times 5} \sum_{m=-2}^2 \sum_{n=-2}^2 Grad_y(i+m, j+n) \right)^2 \right| \quad (9)$$

where $img_sub(i+m, j+n)$ and $img_sub(i, j)$ represent the value of a pixel in a 5×5 rectangle with the coordinate (i, j) as the center and the pixel value at the coordinate (i, j) , respectively. $Grad_x(i+m, j+n)$ and $Grad_y(i+m, j+n)$ are the gradients in the x and y directions, respectively. We used the Scharr operator to calculate the gradient value.

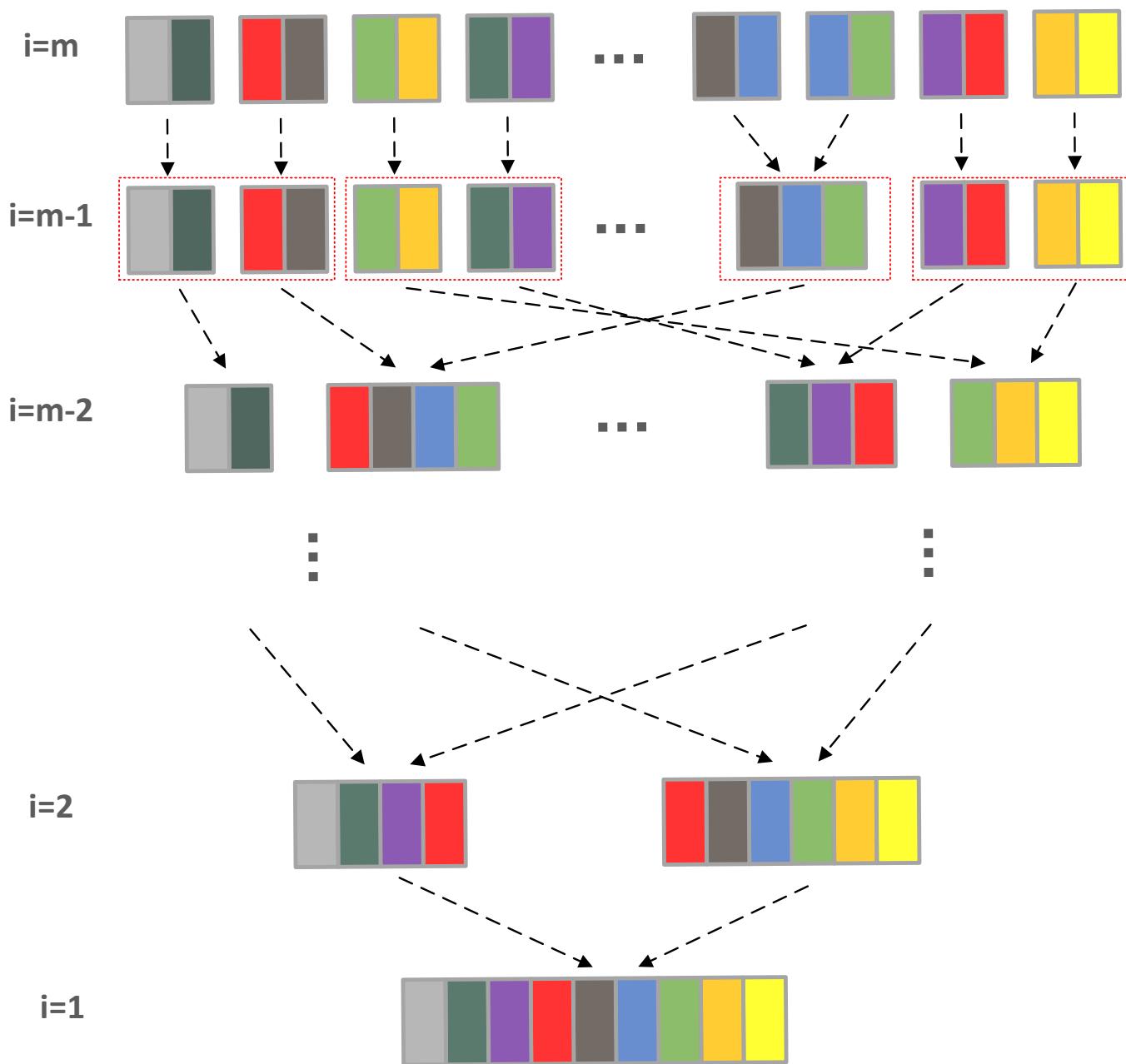


Figure 2. Image mosaic structure diagram of merge sorting. The m -th layer represents the input unordered images, and the first layer represents the final stitched image. The parts with the same colors represent the overlapping areas of the images.

Using dynamic programming to calculate the energy value of each matching point in the overlapping area of the image according to the energy formula (Equation (7)), we were able to find the lowest energy seam from the top to the bottom of the image. We set the pixel with the lowest energy value among the matching points of the first row of the overlapping area of the image as the vertex of the image stitching line, denoting it as $p(i,j)$, and then denoted the three matching points of the next row as $p(i-1, j+1)$, $p(i, j+1)$, and $p(i+1, j+1)$ and calculated their energy values. The j direction of the point with the minimum energy value was the expansion direction of the next row. The matching points were selected in turn downward. When the matching points of the last row were obtained, all the matching points were connected together, forming the optimal splicing line. The N images were set as N nodes in the merge-sorting model, and then

each adjacent pair of images were merged together for feature point matching. When the images met the condition of having the minimum number of matching points, we used dynamic programming to determine the seam line with the lowest energy in the images for fusion and then used them as the upper-layer-node images, while those that did not meet the condition were directly used as the upper-layer-node images. The images in the upper-layer node were stitched according to the steps of Algorithm 1 and, finally, the images of the same scene were stitched into a complete image. The flowchart for the MSIS method is shown in Figure 3. We used unordered images of two scenes, scene 1 and scene 2. The four images of scene 1 were obtained from [28], and the five images of scene 2 were taken with Samsung mobile phones. Using the stitching model based on the merge-sorting algorithm proposed in this paper, the images were first divided into five groups and the feature points of the images were matched when each pair of images was grouped and the stitching conditions were met directly. Then, stitching was performed sequentially. It can be seen that the stitching results for scene 1 and scene 2 differed. The stitching result for scene 1 was a wide-angle image, while the stitching result for scene 2 was a 180° semi-panoramic image. The stitching results are shown in Figure 4.

Algorithm 1 Multi-scene image stitching based on merge sorting

Input: n unordered images.

Output: one or more panoramic images.

1. n images are input as the initial images of the m -th layer and image $[i]$ is set to represent the i -th image of each layer ($i = 1, 2, 3, \dots, n$).
 2. The n images are divided into $n/2$ groups, and each pair of adjacent two images is a group. At the same time, the feature points of the two images in these groups are detected and matched. If the minimum number of matching points is met, image stitching is carried out directly; otherwise, go to step 3.
 3. The first image is taken as the target image, and feature point detection and matching are carried out with the i -th image. If the number of matched feature points reaches the minimum, proceed to step 4; otherwise, proceed to $i + 1$ and continue with step 3.
 4. The images $[i]$ and $[i + 1]$ that meet the requirements are fused using the optimal seam line method to obtain image $[j]$ ($j < n$) and, at the same time, the fused images $[i]$ and $[i + 1]$ from the input image dataset are deleted. If $n = 0$ or $n = 1$, go to step 5; otherwise, go back to step 3.
 5. Taking the image output from step 4 as the input image for the next layer, the number of input images in the next layer is image $[n] = \text{image } [j]$ ($n = j$). When $j = 1$, go to step 6; otherwise, go back to step 3.
 6. The output image of the last layer is the image we want to obtain.
-

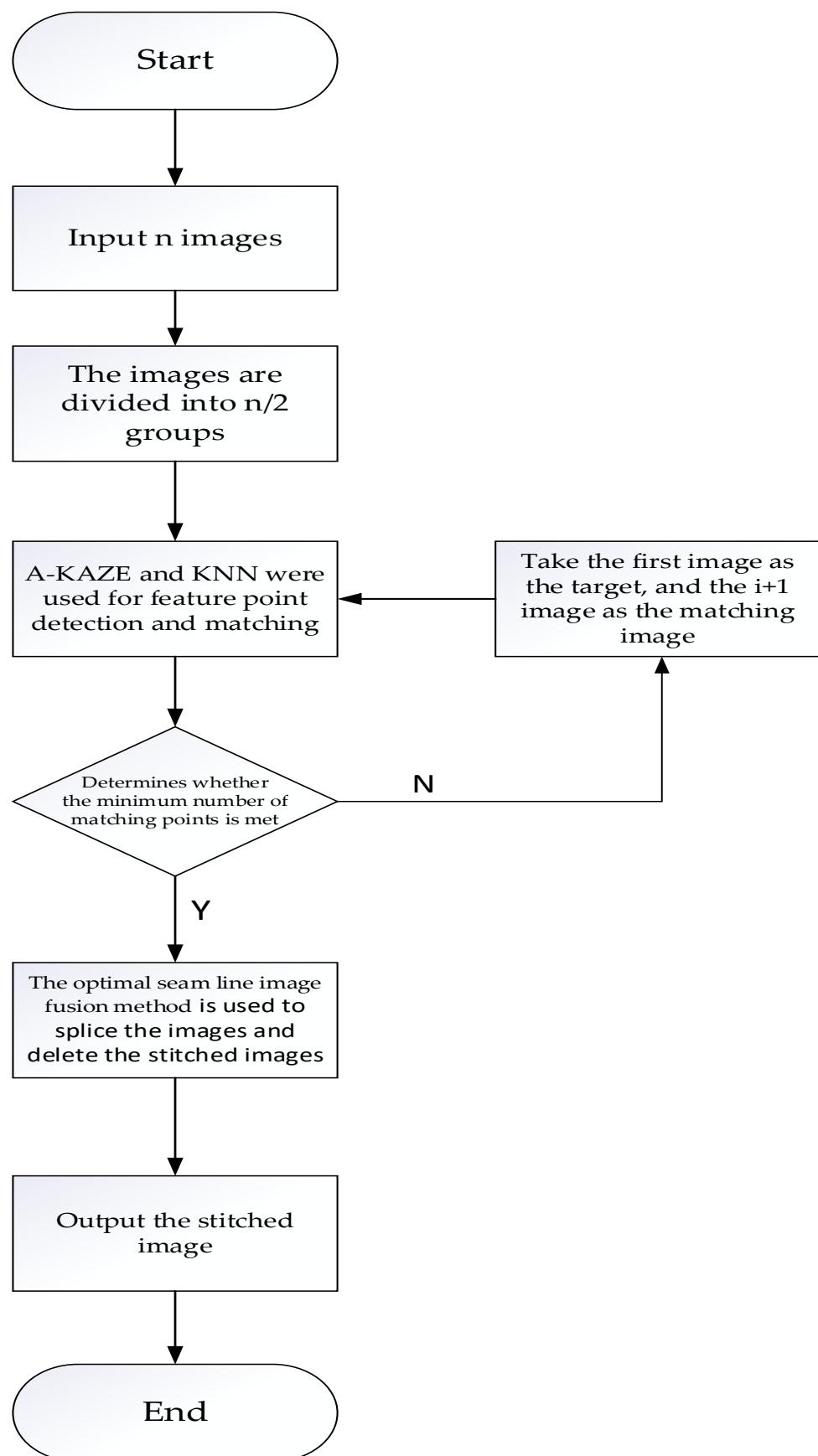


Figure 3. The flowchart of the MSIS method.

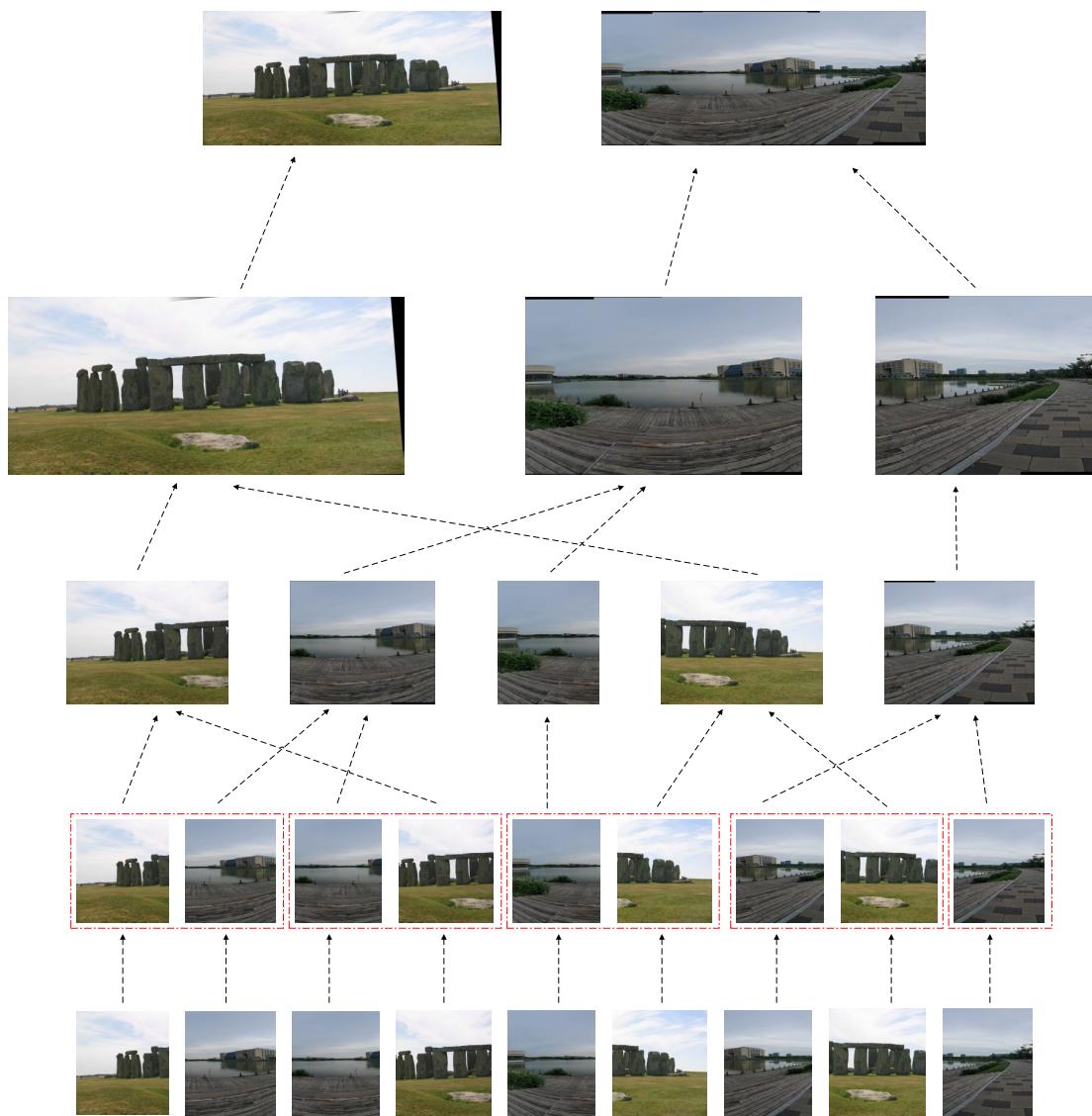


Figure 4. Stitching of unordered images of two scenes. Scene 1 consists of four images; scene 2 consists of five images.

5. Experimental Results and Analysis

In this section, we describe the experiment and the analysis of multiple scenes. The environment used in our experiments employed OpenCV-3.4.2 on a 2.30 GHz CPU with an Intel(R) Core (TM) i5-6300HQ processor and Windows 10 operating system. The image data used in the experiments described in this section were obtained manually using the camera of a Samsung Galaxy S10+ mobile phone and a phone holder.

We compared the matching time required for the feature points of images ab and cd shown in Figure 1. In order to obtain more suitable data, we repeated the experiment five times using the same images with each method and then took the average of these five experiments as the comparison data, as shown in Table 1. It can be seen from Table 1 that the time spent computing and matching the feature points with the three methods increased with increasing image pixels. For image ab, SURF took the most time among the three algorithms, followed by SIFT, while A-KAZE required the least amount of time, about 50% that of SIFT and 20% that of SURF. For image cd, SIFT took the most time, followed by SURF. A-KAZE still required the least amount of time, about 38% that of SIFT and 43% that of SURF, demonstrating an improvement in the efficiency of image feature point matching.

Table 1. Comparison of time taken for feature point computation and matching (unit: ms).

Image	Number of Experiments	SIFT	SURF	A-KAZE
(a) and (b) 800 × 600	1	510	1141	247
	2	551	1122	260
	3	517	1224	252
	4	529	1118	275
	5	501	1132	254
	Average	521	1147	257
(c) and (d) 2592 × 1944	1	7129	6335	2840
	2	7170	6282	2689
	3	7209	6326	2713
	4	7231	6329	2699
	5	7238	6304	2683
	Average	7195	6315	2724

We used two scenes (ordered and unordered images) to verify that the MSIS method worked. Scene 1 comprised 8 (1280 × 2862) ordered images Figures 5–8 show the results for the stitching of scene 1 with the three methods., Scene 2 comprised 14 (1280 × 1556) unordered images Figures 9–12 show the results of the stitching of scene 2 with the three methods. The method based on the binary tree model demonstrated obvious stitching lines in the result for the orderly image stitching with scene 1, and the overlapping areas on both sides of the stitching lines were not well fused together, resulting in a poor stitching effect. For scene 2, in the final stitching process, the image was obviously compressed towards the middle, and the final boundary was distorted, causing the boundary information to be blurred. For the stitching method with the left-to-right approach, the stitched image for scene 1 had stitching line shadows and the final boundary had serious distortion shadow phenomena, resulting in loss of boundary information. For scene 2, there was a serious black shadow at the final boundary of the stitched image, resulting in loss of boundary information and an incompletely stitched image. The MSIS method combines the merge sorting-based model and the optimal seam line method, making it possible to address the phenomenon of stitching line shadows in the stitched image and weaken the distortion of the final boundary in scene 1, reducing the loss of boundary information. In scene 2, the MSIS method was also able to address the phenomenon of distortion at the last boundary of the image well, and it preserved the boundary information and restored the real scene well.

With regard to image stitching efficiency, we repeated the experiment five times using the same image data for each scene and then took the average value for the five experiments as the comparison data, as shown in Table 2. For the ordered image stitching with scene 1, the time required by the method proposed in this paper was reduced by about 30% compared to the binary tree-based model, and the time required was only about 5% more than with the left-to-right method but with a greatly improved stitching effect. For the unordered image stitching with scene 2, the method in this paper not only ensured good results but also took the least amount of time among the three methods. For scene 2, the time required by the MSIS method was reduced by about 10% compared to the binary tree-based model and by about 13% compared to the left-to-right stitching method, demonstrating an improvement in the stitching efficiency with unordered images.

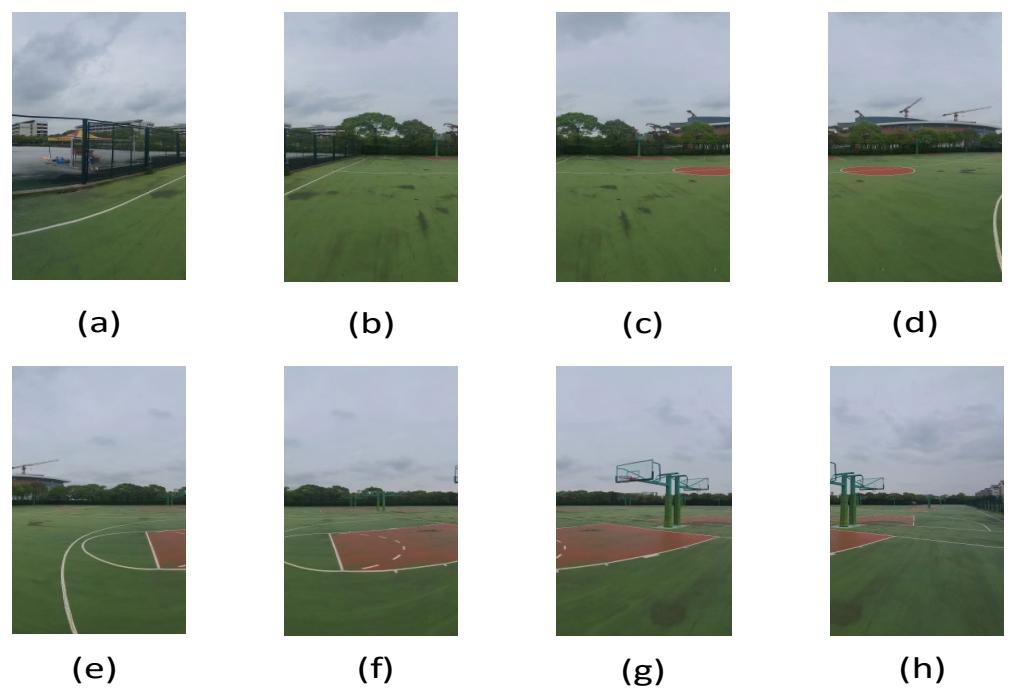


Figure 5. (a–h) are experimental images of scene 1.



Figure 6. The results obtained based on the binary tree model.



Figure 7. Results obtained with the left-to-right approach.



Figure 8. Results obtained with the method proposed in this paper.

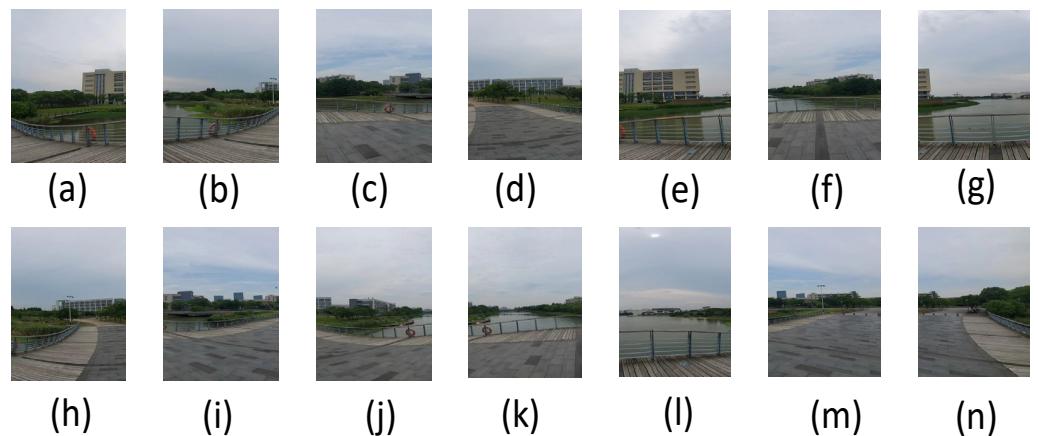


Figure 9. (a–n) are the experimental images of scene 2.



Figure 10. The result of stitching based on the binary tree model.



Figure 11. Result obtained by stitching from left to right.



Figure 12. Result obtained with the MSIS method.

Table 2. Time required for stitching images with each method for the two experimental scenes (unit: s).

Scene	Number of Experiments	Time Spent Stitching Images		
		Binary Tree-Based Model	Left-to-Right Approach	MSIS
Scene 1 (8 images) (1280 × 2565)	1	736	525	626
	2	720	517	512
	3	738	497	502
	4	747	503	513
	5	712	505	518
Average	6	730	509	534
Scene 2 (14 images) (1280 × 1556)	1	734	798	665
	2	725	803	654
	3	730	813	659
	4	718	806	648
	5	720	810	655
Average	6	725	806	656

The above experiments show that the method proposed in this paper could not only achieve the stitching of multi-scene images and produce a good stitching effect but also had improved stitching efficiency.

6. Conclusions

In contrast to the traditional image stitching method, a multi-scene image stitching method based on the merge-sorting model was proposed in this paper. The bottom-up stitching method is adopted for the input image to construct the merge-sorting model, and the root-node image in the merge-sorting model is the final stitched image. For image registration, the A-KAZE features are extracted first, and then the feature points are matched using the bidirectional KNN algorithm. This makes it possible to avoid spending a lot of time with image registration and improves the efficiency of image registration. When there are two images that match well, this method stitches them directly without sorting, which greatly reduces the number of registrations of unrelated images. The experimental results show that the proposed method can improve the efficiency of image stitching, enhance the robustness of image registration, and reduce the distortion in the stitched images. The MSIS method improves the stitching efficiency and reduces the distortion in the images after stitching. However, the image registration takes a long time throughout the process. In addition, stitched images have black boundary bars, and there are performance limitations relating to the device used for processing. In the future, we will continue our research and further reduce the time required for image registration. For image fusion, we will continue to investigate the black bars shown in Figure 7 and upgrade the corresponding processing equipment.

Author Contributions: Methodology, W.L.; writing—original draft, W.L.; software, W.L. and J.H.; supervision, B.S.; writing—review and editing, K.Z., Y.Z. and J.H.; validation, K.Z.; funding acquisition, B.S.; investigation, B.S.; project administration, B.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 32271248.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study.

Acknowledgments: The authors thank the staff from the Experimental Auxiliary System and the Information Center of Shanghai Synchrotron Radiation Facility (SSRF) for on-site assistance with the GPU computing system.

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

References

- Zhang, L.; Chen, J.; Liu, D.; Shen, Y.; Zhao, S. Seamless 3D Surround View with a Novel Burger Model. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; pp. 4150–4154. [CrossRef]
- Alwan, M.G.; Al-Brazinji, S.M.; Mosslah, A.A. Automatic panoramic medical image stitching improvement based on feature-based approach. *Period. Eng. Nat. Sci. (PEN)* **2022**, *10*, 155–163. [CrossRef]
- Zhang, T.; Zhao, R.; Chen, Z. Application of Migration Image Registration Algorithm Based on Improved SURF in Remote Sensing Image Mosaic. *IEEE Access* **2020**, *8*, 163637–163645. [CrossRef]
- Zhang, Y.; Zhou, D.; Xing, Y.; Cen, M. An Adaptive 3D Panoramic Vision System for Intelligent Vehicle. In Proceedings of the 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 22–24 May 2021; pp. 2710–2715. [CrossRef]
- Liao, T.; Li, N. Single-Perspective Warps in Natural Image Stitching. *IEEE Trans. Image Process.* **2019**, *29*, 724–735. [CrossRef] [PubMed]
- Nie, Y.; Su, T.; Zhang, Z.; Sun, H.; Li, G. Dynamic Video Stitching via Shakiness Removing. *IEEE Trans. Image Process.* **2017**, *27*, 164–178. [CrossRef] [PubMed]
- Qu, Z.; Li, J.; Bao, K.-H.; Si, Z.-C. An Unordered Image Stitching Method Based on Binary Tree and Estimated Overlapping Area. *IEEE Trans. Image Process.* **2020**, *29*, 6734–6744. [CrossRef] [PubMed]
- Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]
- Herbert, B.; Andreas, E.; Tinne, T.; Luc, V.G. Speeded-up robust features (SURF). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359. [CrossRef]
- Xu, X.; Tian, L.; Feng, J.; Zhou, J. OSRI: A Rotationally Invariant Binary Descriptor. *IEEE Trans. Image Process.* **2014**, *23*, 2983–2995. [CrossRef]
- Alcantarilla, P.; Nuevo, J.; Bartoli, A. Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces. *IEEE Trans. Patt. Anal. Mach. Intell.* **2011**, *34*, 1281–1298.
- Abu Bakar, S.; Jiang, X.; Gui, X.; Li, G.; Li, Z. Image Stitching for Chest Digital Radiography Using the SIFT and SURF Feature Extraction by RANSAC Algorithm. *J. Physics Conf. Ser.* **2020**, *1624*, 042023. [CrossRef]
- Li, Y.; Wang, Y.; Huang, W.; Zhang, Z. Automatic image stitching using SIFT. In Proceedings of the 2008 International Conference on Audio, Language and Image Processing, Shanghai, China, 7–9 July 2008; pp. 568–571. [CrossRef]
- Wang, W.; Chang, F. A Multi-focus Image Fusion Method Based on Laplacian Pyramid. *J. Comput.* **2011**, *6*, 2559–2566. [CrossRef]
- Cao, Q.; Shi, Z.; Wang, P.; Gao, Y. A Seamless Image-Stitching Method Based on Human Visual Discrimination and Attention. *Appl. Sci.* **2020**, *10*, 1462. [CrossRef]
- Avidan, S.; Shamir, A. Seam carving for content-aware image resizing. *ACM SIGGRAPH* **2007**, *26*, 10. [CrossRef]
- An, Q.; Chen, X.; Wu, S. A Novel Fast Image Stitching Method Based on the Combination of SURF and Cell. *Complexity* **2021**, *2021*, 9995030. [CrossRef]
- Dong, S.; Lu, L. Transmission line image mosaic based on improved SIFT-LATCH. *J. Physics: Conf. Ser.* **2021**, *1871*, 012060. [CrossRef]
- Zhu, J.T.; Gong, C.F.; Zhao, M.X.; Wang, L.; Luo, Y. Image mosaic algorithm based on pca-orb feature matching. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *42*, 83–89. [CrossRef]
- Lu, J.; Huo, G.; Cheng, J. Research on image stitching method based on fuzzy inference. *Multimed. Tools Appl.* **2022**, *81*, 23991–24002. [CrossRef]
- Zhang, L.; Huang, H. Image Stitching with Manifold Optimization. *IEEE Trans. Multimed.* **2022**, *1*. [CrossRef]
- Chen, M.; Zhao, X.; Xu, D. Image Stitching and Blending of Dunhuang Murals Based on Image Pyramid. *J. Phys. Conf. Ser.* **2019**, *1335*, 012024. [CrossRef]
- Snavely, N.; Seitz, S.M.; Szeliski, R. Modeling the World from Internet Photo Collections. *Int. J. Comput. Vis.* **2008**, *80*, 189–210. [CrossRef]
- Bay, H.; Tuytelaars, T.; Van Gool, L. Surf: Speeded up robust features. *Eur. Conf. Comput. Vis.* **2006**, *3951*, 404–417. [CrossRef]
- Wei, C.; Yu, L.; Yawei, W.; Jing, S.; Ting, J.; Qinglin, Z. Fast image stitching algorithm based on improved FAST-SURF. *J. Appl. Opt.* **2021**, *42*, 636–642. [CrossRef]
- Tareen, S.A.K.; Saleem, Z. A comparative analysis of sift, surf, kaze, akaze, orb, and brisk. In Proceedings of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 3–4 March 2018; pp. 1–10. [CrossRef]
- Ma, H.; Jia, C.; Liu, S. Multisource image fusion based on wavelet transform. *Int. J. Inf. Technol.* **2005**, *11*, 81–91.

28. Lytt-2020, Image Data for Image-Stitching Project. 2022. Available online: <https://github.com/Lytt-2020/images-stitching> (accessed on 20 August 2022).
29. Qu, Z.; Bu, W.; Liu, L. The algorithm of seamless image mosaic based on A-KAZE features extraction and reducing the inclination of image. *IEEE Trans. Electr. Electron. Eng.* **2018**, *13*, 134–146. [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.