



Article Optimized Seam-Driven Image Stitching Method Based on Scene Depth Information

Xin Chen, Mei Yu * and Yang Song

Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China; wzxc07@163.com (X.C.); songyang@nbu.edu.cn (Y.S.) * Correspondence: yumei@nbu.edu.cn

Abstract: It is quite challenging to stitch images with continuous depth changes and complex textures. To solve this problem, we propose an optimized seam-driven image stitching method considering depth, color, and texture information of the scene. Specifically, we design a new energy function to reduce the structural distortion near the seam and improve the invisibility of the seam. By additionally introducing depth information into the smoothing term of energy function, the seam is guided to pass through the continuous regions of the image with high similarity. The experimental results show that benefiting from the new defined energy function, the proposed method can find the seam that adapts to the depth of the scene, and effectively avoid the seam from passing through the salient objects, so that high-quality stitching results can be achieved. The comparison with the representative image stitching methods proves the effectiveness and generalization of the proposed method.

Keywords: image stitching; image depth; seam-cutting



Citation: Chen, X.; Yu, M.; Song, Y. Optimized Seam-Driven Image Stitching Method Based on Scene Depth Information. *Electronics* **2022**, *11*, 1876. https://doi.org/10.3390/ electronics11121876

Academic Editor: Gwanggil Jeon

Received: 21 April 2022 Accepted: 12 June 2022 Published: 14 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

Image stitching [1] refers to the use of a set of images of the same scene taken from different perspectives to create a single fused image with a wider field of view. It is widely used in multimedia content generation, image analysis/understanding, industrial inspection, and other fields (such as panoramic imaging [2], aerial image generation [3], medical synthetic image generation [4,5], virtual reality [6], remote visual inspection [7], and so on).

Image stitching methods have been applied to more and more scenarios. In order to adapt to these scenarios, many methods have been proposed or improved. Currently, image stitching is mainly implemented in two ways: the spatially varying warping method and the seam-driven method.

For the first way, Brown et al. [8] carried out image stitching with spatially varying warping, which aligns the input images by estimating the optimal homography matrix of the entire images. When the input images have only translation or rotation transformation, or the image scene is coplanar, this type of method can obtain visually acceptable stitching results. However, such a method may introduce visual artifacts or misalignment to the stitched image when the input images are of large parallax. In order to alleviate the problems caused by parallax, Gao et al. [9] proposed a dual homography matrix model to stitch images, and assumed that the scene can be modeled by two depth planes; however, when the scene has more than two depth planes, the quality of stitched image will decrease. To solve the adaptability of the method to small parallax scenes, Zaragoza et al. [10] proposed an as-projective-as-possible image stitching (APAP) model to improve the accuracy of image alignment and reduce the ghosting effect through flexible local deformation. Although the local deformation method provides accurate alignment, serious perspective distortion will occur in the non-overlap regions. Therefore, Lin et al. [11] proposed the adaptive as-natural-as-possible (AANAP) warping method for the images with unnatural rotations. Li et al. [12] proposed a parallax-tolerant image stitching method based on robust elastic

warping (REW). Liao et al. [13] designed single-perspective warps (SPWs) for natural image stitching, which introduces point and line features to further improve the naturalness of the stitched image while ensuring image alignment. Shi et al. [14] pointed out that the warping methods used to regard stitching as the construction of a geometric transformation model, which limits the deformation effect in areas where the depth of field changes dramatically, and few studies focused on post-processing to further eliminate these projection biases after image warping. For this reason, they proposed a misalignment-eliminated warping image stitching based on grid-based motion statistics matching.

Different from spatially varying warping, the seam-driven method tries to find a seam in the overlapping region of the aligned images. In the stitched image, the contents on both sides of the seam come from different aliened images, respectively. This type of method does not need to strictly align the entire overlapping region, but only the regions near the seam. Therefore, the seam-driven image stitching method can handle the parallax problem well if a suitable seam is found, which means that the search for a suitable seam is crucial to the stitching result. Gao et al. [15] first proposed a seam-driven image stitching method, and defined a seam quality metric to measure the effectiveness of the seam. Huang et al. [16] proposed a seam planning method, aiming to maximally preserve the visual content while eliminating inconsistency in the overlapping region. Lin et al. [17] proposed a seam-guided local alignment (denoted as SEAGULL) method for parallax-tolerant image stitching, in which the final stabilized warp is accomplished through iteratively computing the seam location and the structure-preserving warping. Li et al. [18] proposed a perception-based seam cutting method for image stitching which takes into account the non-linearity and non-uniformity of human perception in energy minimization, and obtained a substantial improvement over the traditional seam-cutting approach. Herrmann et al. [19] used an object detection technique to extract the center of the object, and modified the energy function at the seam search stage to improve the anti-occlusion ability of the method. Wang et al. [20] used curve transformation to detect seam in images to be stitched and improve the stitching quality of images. Generally, for overlapping region of given aligned images, different energy functions will result in different seams, and finally lead to different stitching results. Therefore, in order to obtain reasonable image stitching results, it is necessary to find an suitable seam to complete the image stitching.

Although the above methods have achieved good results in solving the misalignment caused by parallax, it is still a challenge when the scene has continuous depth changes and the image texture is complex. To solve this problem, we propose an optimized seam-driven image stitching method considering depth, color, and texture information of the scene. Specifically, in view of the structural distortion of the existing seam-driven image stitching methods, we design a new energy function to reduce the structural distortion near the seam and improve the invisibility of the seam. By additionally introducing depth information into the smoothing term of energy function, the seam is guided to pass through the continuous regions of the image with high similarity, so as to avoid the seam passing through the protruding objects. Experimental results show that by using the improved energy function additionally integrating depth information, the proposed seam-driven image stitching method can effectively deal with large parallax scenes with continuous depth changes.

The remainder of this paper is organized as follows. Section 2 presents the motivation of this work and describes the proposed method in detail. Section 3 provides the experimental results and analysis, and the effectiveness of the proposed method is additionally proved through objective evaluation indicator. Finally, the conclusion is given in Section 4.

2. Seam-Driven Image Stitching Based on Depth, Color, and Texture Information

In this section, we first describe the principle of the seam-driven image stitching method and the motivation of this work. Then, we propose an optimized seam-driven image stitching method based on depth, color, and texture information of the scene. Different from the existing methods, the proposed method additionally considers the scene depth information in the energy function for seam search, so as to guide the seam bypassing the protruding objects and improve the invisibility of the seam.

2.1. Motivation

Different from the spatially varying warping methods, the seam-driven image stitching methods can handle the parallax problems well in many cases, and they can avoid misalignment by finding suitable seam in regions with simple or well-arranged textures. This type of method does not require strict alignment of the entire overlapping region, but only the region near the seam.

Taking the stitching of two images as an example, let I_0 and I_1 be a pair of aligned images, **P** represent the overlapping region of I_0 and I_1 , and **L** = {0,1} be a label set, where "0" and "1" are related to I_0 and I_1 , respectively. Then, the seam search problem can be described as a segmentation problem, which is equivalent to a binary label problem. The process of seam-driven image stitching is to assign a label $l_p \in \mathbf{L}$ to each pixel $p \in \mathbf{P}$. $l_p = 0$ means that the value of the pixel p should be copied from the image labeled with "0", while $l_p = 1$ indicates that the value of the pixel p will be copied from the image labeled with "1". The goal of the seam cutting is to find suitable label l that minimizes the energy function E(l).

$$E(l) = \sum_{p \in \mathbf{P}} E_d(p, l_p) + \sum_{(p,q) \in \mathbf{N}} E_s(l_p, l_q)$$
(1)

where **N** is the set of pixels in the overlapping region, the data term $E_d(p, l_p)$ is the cost of assigning a label l_p to the pixel p ($p \in \mathbf{P}$), and the smoothing term $E_s(l_p, l_q)$ is the cost of assigning a label (l_p, l_q) to a pair of pixels (p, q) $\in \mathbf{N}$.

The data item $E_d(p, l_p)$ treats the pixels in the overlapping region equally, and penalizes the pixels in the non-overlapping region, so that the seam can only fall in the overlapping region of the images. According to the formula in [21], the data item $E_d(p, l_p)$ can be computed as

$$E_d(p, l_p) = \begin{cases} 0, & \text{if } p \in \text{overlaping region} \\ \lambda_m, & \text{otherwise} \end{cases}$$
(2)

where λ_m represents a large penalty.

The smoothing term $E_s(l_p, l_q)$ represents the cost of discontinuity between pixel *p* and other pixels in its neighborhood. The smaller the difference between the pixels *p* and *q*, the better the invisibility of the seam passing through them, and hence the smaller the corresponding cost should be. $E_s(l_p, l_q)$ can be computed as

$$E_s(l_p, l_q) = \frac{1}{2} |l_p - l_q| \cdot (I_*(p) + I_*(q))$$
(3)

$$I_*(\cdot) = \|I_0(\cdot) - I_1(\cdot)\|_2$$
(4)

where $I_*(\cdot)$ denotes the Euclidean metric difference.

Through the graph cut optimization algorithm [22], the energy function in Equation (1) is minimized and the seam corresponding to the minimal cost can be determined. Usually, the color and texture differences are taken into account in the energy function. However, in the face of consumer-level shooting environments, the depth of the scene often changes greatly, and the texture of the overlapping region is complex and changeable. In this case, it is not enough only to consider the color and texture difference in the energy function.

Figure 1 shows two images to be stitched, and the stitching results obtained with different methods. It is seen that there are regions of continuous depth variation in the pair of images. As shown in Figure 1b, the spatially varying warping method in [13] achieves good alignment at the background where the parallax changes are not obvious, but misalignment occurs in the regions of traffic separation columns where the parallax changes greatly, resulting in artifacts in this region. However, for the perception-based seam-driven method in [18], even though it can avoid most artifacts in the overlapping region, structural distortion still appears in the overlapping region with continuous depth



changes, leading to scene fracture [23] at the first traffic separation column, as shown in Figure 1c.

Figure 1. Failure cases of two types of stitching method. (a) Two images to be stitched. (b) Image stitching result of the method in [13]. (c) Image stitching result of the method in [18]. (d) Image stitching result of the proposed method.

Therefore, an energy function that integrates depth, texture, and color information of the scene is proposed in this paper, so that the found seam can achieve the image stitching results adapting to the depth of the scene, consistent with the human eye perception. Figure 1d shows relatively better stitching result using the method proposed in this paper, since the used seam bypasses the traffic separation columns with continuous depth change. However, Figure 1d still shows unnatural blending along the curb of the paved road. The reason is that the brightness of the two images to be stitched is inconsistent, as shown in Figure 1a. This kind of artifact can be reduced by pre-processing such as brightness correction of the images to be stitched.

2.2. The Proposed Method

In this subsection, we propose an optimized seam-driven image stitching method that additionally integrates depth information of the scene in the energy function, and its framework is shown in Figure 2. Firstly, the input images are pre-aligned through image alignment, and then depth estimation and texture map generation are performed to obtain the depth value and texture feature of the overlapping region of the images to be stitched, based on which the energy function of the overlapping region can be calculated. After that, the graph cut optimization algorithm is used to obtain the seam through minimizing the energy function. Finally, the aligned images are fused with the Poisson fusion method to generate the final stitched image.



Figure 2. The framework of the proposed method.

For the two images to be stitched, let I_L and I_R denote the left and right images, respectively. In this paper, we use the SPW method [13] to pre-align the input images, because it has good alignment capabilities, while ensuring the naturalness of the non-overlapping region of the image, reducing projection distortion, and maintaining strong flexibility and robustness. Through the SPW method [13], two aligned images can be obtained in the same coordinate system, denoted as I_0 and I_1 . It should be noted that even

though the I_L and I_R are aligned, simply fusing I_0 and I_1 may result in ghost or artifacts in the overlapping region of the fused image due to misalignment and moving objects. It is a good alternative to find a suitable seam in the overlapping region of the two aligned images and then copy parts of the two images to both sides of the seam, respectively. Of course, in this way, it is expected that the image contents on both sides of the seam are well stitched, i.e., the seam should be invisible as much as possible.

Obviously, energy function is critical to finding the suitable seam. The early methods usually consider color feature of images in the energy function, but ignore the salience of the object. Therefore, more complex features are included in difference cost of the energy function, and the most commonly used is the combination of color and gradient features, which tries to enhance color consistency while avoiding the seam passing through prominent objects. However, when the scene depth changes continuously, the method combining color and gradient features may have the problem of finding too single seam—for example, the seam found is approximately a straight line and does not bypass the prominent foreground, leading to structural distortion in the stitched image [23]. A perceptually pleasing seam should be usually along some specific regions [24], such as roads, woodlands, and the sky, and it is hoped that the seam can bypass prominent foreground objects. Therefore, a new difference cost is proposed in this paper so as to measure the similarity between the overlapping regions of I_0 and I_1 . Since depth information is usually well associated with objects in the scene, in addition to the color and texture features, we also introduce the depth information into the energy function to improve the understanding of the scene and avoid the signal discontinuity on both sides of the seam as much as possible.

In this work, the advanced monocular depth estimation method in [25] is used to obtain the depth map of the images to be stitched, and in order to obtain reliable texture structure, the method in [26] is utilized to extract the texture map of the images to be stitched. Figure 3 shows the examples of the obtained depth maps and texture maps of the input images, where the texture map is shown in the form of a heat map for clear display.







(b)



(c)

Figure 3. Depth maps and texture maps of different scenes. (a) Testing scenes. (b) Depth maps obtained with the method in [25]. (c) Texture maps obtained with the method in [26] where heat maps are used for clear display.

Let D_0 , D_1 , T_0 , and T_1 denote the depth map and texture map of the images I_0 and I_1 respectively. Then, we define a new difference cost *S* of the overlapping region of the two images to be stitched, which combines color, texture, and depth information of the scene, and is expressed as follows

$$S(\cdot) = \|C_{color}(\cdot)\|_{2} + \|C_{texture}(\cdot)\|_{2} + \|C_{depth}(\cdot)\|_{2}$$
(5)

where $C_{color}(\cdot)$ represents the cost term of color difference, $C_{texture}(\cdot)$ is the cost term of texture difference, and $C_{depth}(\cdot)$ denotes the cost term of texture–color difference combined with depth information.

The cost term of color difference $C_{color}(\cdot)$ is computed by

$$C_{color}(\cdot) = I_0(\cdot) - I_1(\cdot) \tag{6}$$

where $I_0(\cdot)$ and $I_1(\cdot)$ represent the color of pixels corresponding to each other in the overlapping region of I_0 and I_1 . The cost term of color difference is used to guide the seam to pass through regions with similar color as much as possible so as to hide the seam.

The cost term of texture difference $C_{texture}(\cdot)$ is computed by

$$C_{texture}(\cdot) = T_0(\cdot) - T_1(\cdot) \tag{7}$$

where $T_0(\cdot)$ and $T_1(\cdot)$ are the texture value of pixels corresponding to each other in the overlapping region of $T_0(\cdot)$ and $T_1(\cdot)$. The cost term of texture difference makes the seam bypass the regions with complex texture which are prone to structural distortion.

The cost term of the texture–color difference combined with depth information $C_{depth}(\cdot)$ is defined as

$$C_{depth}(\cdot) = T_0(\cdot)D_0(\cdot)e^{-T_0(\cdot)I_0(\cdot)} - T_1(\cdot)D_1(\cdot)e^{-T_1(\cdot)I_1(\cdot)}$$
(8)

where $D_0(\cdot)$ and $D_1(\cdot)$ denote the depth value of pixels corresponding to each other in the overlapping region of D_0 and D_1 , respectively. Due to the integration of depth information, $C_{depth}(\cdot)$ can optimize the seam-driven image stitching method to keep away from the region of significant depth change at object boundaries, so that the obtained seam can better adapt to image stitching of scenes with large parallax or scenes with continuous parallax changes.

Thus, Equation (3) can be rewritten as

$$\overline{E}_{s}(l_{p}, l_{q}) = \frac{1}{2} |l_{p} - l_{q}| (S_{*}(p) + S_{*}(q))$$
(9)

The final energy function is defined by

$$\overline{E}(l) = \sum_{p \in \mathbf{P}} E_d(p, l_p) + \sum_{(p,q) \in \mathbf{N}} \overline{E}_s(l_p, l_q)$$
(10)

The seam can be obtained through minimizing the energy function E(l) with the graph cut optimization algorithm [22], and then the aligned images can be fused with the Poisson fusion strategy [27] to obtain the final stitched image.

3. Experimental Results and Analysis

At present, the image stitching is mainly implemented in two different ways [17]: the spatially varying warping method and the seam-driven method. Therefore, the proposed method will be compared with these two types of representative methods in this section. For the sake of fairness, these comparative experiments are conducted on the two public datasets [17,28].

3.1. Comparison with Spatially Varying Warping Methods

In order to demonstrate the effectiveness of the proposed seam-driven method for large parallax scenes, it is firstly compared with four spatially varying warping methods, namely APAP [10], AANAP [11], REW [12], and SPW [13].

The pair of input images shown in Figure 4a,b comes from the literature [17]. It is seen that there are prominent foreground and objects with continuous parallax changes in the input images; hence, they are used to evaluate the effectiveness of the methods. Figure 4d–g show the stitching results of APAP [10], AANAP [11], REW [12], and SPW [13] methods, respectively. As foreground objects, the close-range objects in the red and green boxes have large parallax, which can be used to test the different methods' ability to process objects with large parallax. From the partial enlargement in Figure 4d–g, it can be seen that there are obvious ghosts in the results of APAP [10], AANAP [11], REW [12], and SPW [13] methods. By contrast, Figure 4c shows the difference cost of the overlapping region calculated with the proposed method, and Figure 4i shows the seam found under the guidance of the proposed energy function. Since the seam bypasses the prominent foreground objects, there is no artifact in the stitched image obtained with the proposed method, as shown in Figure 4h, indicating that the proposed method can effectively process objects with large parallax.



(a)

(b)

(c)



(**d**)







(e)



Figure 4. The stitching results of different methods. (a,b) Two input images to be stitched. (c) Difference cost of the overlapping region obtained with the proposed method, where heat map is

used for clear display. (d) Results of APAP [10]. (e) Results of AANAP [11]. (f) Results of REW [12]. (g) Results of SPW [13]. (h) Results of the proposed method. (i) The seam obtained with the proposed method.

3.2. Comparison with Seam-Driven Methods

In order to verify the rationality of the proposed method for introducing depth information into the difference cost in seam search, it is also compared with the image stitching method based on perceptual seam-cutting [18] and the traditional seam-driven method based on color and texture information. In order to ensure the fairness of the comparison, the alignment approach used in the SPW method [13] is utilized to pre-align the input images for all comparison methods. Since the dataset in [17] only provides the final stitched image, the final results in [17] are used as the benchmark in the comparison experiment, focusing on comparing the locally enlarged details of the corresponding regions. Moreover, as the regions through which the seam passes may have structural distortion, we will focus on comparing the final stitching results of such regions.

The perceptual seam-cutting method in [18] takes into account the non-linearity and non-uniformity of human eye perception in the energy minimization. Compared with the traditional seam-driven method based on color and texture information, there is great improvement, but still has some problems. From the results in the second row of Figure 5a, it can be found that the structural distortion still appears at the top of the pool fence where the seam passes through, since the depth of this region (water and pool fence) changes significantly, compared with the partial enlargement of the result obtained with the method in [17], as shown in the third row of Figure 5a. In addition, because the perceptual seam-cutting method in [18] only considers the perception of color discrimination and salient objects, the used seam is more inclined to pass through the region where the image color change is not obvious, leading to the white clouds in the sky above the tallest building being cut off in the final stitched image, as shown in the first row of Figure 5a.



Figure 5. Comparison of three seam-driven methods and the used seams. (a) The perceptual seamcutting method [18]. (b) The traditional seam-driven method based on color and texture information. (c) The proposed method. The first row shows the stitching results and the used seams with respective

to the three different methods. The second row shows partial enlargements of the region through which the seam passes. The third row shows the partial enlargements of the region in the final stitched image of the SEAGULL dataset [17] which correspond to that shown in the second row. The third row is used as the comparison for the second row.

Figure 5b shows the image stitching result of the traditional seam-driven method based on color and texture information, in which the seam tends to pass through regions with similar colors. Due to the constraint of texture information, the seam will bypass the regions with complex texture. However, because of the lack of understanding of the depth of the scene, the finally found seam directly passes through the boundary of the overlapping region of the image (see the first row of Figure 5b and the input image shown in Figure 4b), causing serious structural distortion in the stitched image. Compared with the third row of Figure 5b obtained with the method in [17], it can be seen that the chair legs in the foreground are malposed, and the pool fence is also broken, as shown in the second row of Figure 5b.

Figure 5c is the result of the proposed method. Since the color, texture, and depth information of the scene are comprehensively considered, the understanding ability of the stitching method to the scene is prompted. Therefore, the proposed method can find the seam according to the depth of the scene and bypass the prominent foreground regions. As shown in the second row of Figure 5c, the proposed method does not produce structural distortion in the edge regions where the depth changes significantly. By contrast, the result of the method in [17] has a stitching error. There are two duplicated black dots appearing in the center right of the stitched image, as shown in the third row of Figure 5c.

Figures 6 and 7 give an example of image stitching where the parallax between the two input images is extremely large. The two input images to be stitched and the corresponding difference cost map calculated by the proposed method are given in Figure 6. Figure 7 shows the comparison of the stitching results of the scene in Figure 6. Compared with the partial enlargement of the region in the final stitched image in dataset [17] shown in the third row of Figure 7a, the result of the perceptual seam-cutting method in [18] has serious structural distortion at the building, as shown in the second row of Figure 7a, because the seam passes through this background building. However, for the traditional seam-driven method based on color and texture information, as shown in the second row of Figure 7b, there are obvious stitching errors, the upper part of the green stick disappears, and the tower crane behind also shows geometric distortion. By contrast, benefiting from the introduction of scene depth information in the energy function, the seam found by the proposed method bypasses the building and tower crane at the background, avoiding the structural distortions, as shown in Figure 7a,b, and the pontoon at the foreground is also well stitched, as shown in the second row of Figure 7c.



(a)

(b)

(c)

Figure 6. Input images and their difference cost map. (**a**,**b**) Two input images to be stitched. (**c**) Difference cost map of the overlapping region obtained with the proposed method, where the heat map is used for clear display.



(a)

(b)

(c)

Figure 7. Comparison of three seam-driven methods and the used seams. (a) The perceptual seamcutting method [18]. (b) The traditional seam-driven method based on color and texture information. (c) The proposed method. The first row shows the stitching results and the used seams with respective to the three different methods. The second row shows partial enlargements of the region through which the seam passes. The third row shows the partial enlargements of the region in the final stitched image of the SEAGULL dataset [17] which correspond to that shown in the second row. The third row is used as the comparison for the second row.

3.3. Objective Quality Evaluation

In seam-driven image stitching, an unsuitable seam will produce visual artifacts, which is resulted from the structure inconsistency between the two side of the seam [23]. Here, a seam quality metric is used to quantitatively measure the effectiveness of a seam. Specifically, for each pixel p_i on a seam, a 15 × 15 local patch centered at p_i can be determined. Then, let S_{ZNCC} denote the zero-normalized cross-correlation score between the local patch in the left image and the corresponding patch in the right image. Then, it can be calculated as follows

$$S_{ZNCC} = \frac{1}{n} \sum_{x,y} \frac{1}{\sigma_L \sigma_R} (L(x,y) - \mu_L) (R(x,y) - \mu_R)$$
(11)

where L(x,y) and R(x,y) represent the pre-aligned left and right images, i.e., I_0 and I_1 , respectively. *n* is the number of pixels in the local patches. μ_L and μ_R are the mean value of local patches in the pre-aligned left and right images, while σ_L and σ_R denote the variance of local patches in the pre-aligned left and right images, respectively.

Finally, along the seam, the quality of the seam is calculated by

$$Q = \frac{1}{m} \sum_{i=1}^{m} \left(1 - \frac{S_{ZNCC}(p_i) + 1}{2} \right)$$
(12)

where m is the total number of pixels on the seam. The smaller the quality score Q is, the more reasonable the seam position is, and the better the final stitching effect is.

Here, Equation (12) is used to quantitatively measure the performance of the perceptual seam-cutting method in [18], as well as the traditional seam-driven method based on color and texture information and the proposed method. There are totally 24 sets of test images, all of which are from the public datasets [17,28]. Table 1 shows the quality score *Q* of the proposed method in comparison with the other two methods, where the best results are in bold.

Scene No.	01.	02.	03.	04.	05.	06.	07.	08.	09.	10.	11.	12.
Perceptual [18]	0.121	0.356	0.338	0.378	0.253	0.445	0.290	0.376	0.354	0.360	0.418	0.351
Traditional	0.236	0.527	0.366	0.435	0.332	0.433	0.352	0.320	0.426	0.353	0.409	0.318
Proposed	0.082	0.308	0.346	0.281	0.263	0.421	0.258	0.273	0.352	0.345	0.321	0.239
Scene No.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.
Perceptual [18]	0.318	0.232	0.401	0.360	0.330	0.383	0.248	0.294	0.341	0.306	0.397	0.212
Traditional	0.305	0.262	0.379	0.387	0.355	0.418	0.259	0.434	0.445	0.287	0.348	0.312
Proposed	0.232	0.213	0.371	0.345	0.287	0.389	0.195	0.308	0.431	0.251	0.334	0.276

Table 1. Objective quality scores *Q* of seams obtained with different methods.

From Table 1, it is found that the proposed method is superior to the traditional seamdriven method based on color and texture information and the perceptual seam-cutting method in [18] for most of the test images. The five rows of Figure 8 show the stitched images and the used seams of the 1st, 2nd, 6th, 11th, and 19th scenes in turn, including indoor scenes as well as outdoor scenes with different contents. Figure 8b shows that the proposed method can find the seam according to the depth of the scene to bypass prominent objects, and in the face of a scene with complex texture, it will also search for the seam based on the edge of the object's texture, finally obtaining the seam passing through the regions that are not sensitive to human eyes. In comparison with the results of the SEAGULL dataset [17] shown in Figure 8c, the subjective quality of the stitched images obtained by the proposed method is basically the same, and no obvious structural distortion can be seen by human eyes. In comparison with the results of the seam found by the proposed method is more reasonable because it reduces the situation that the seam extends along the boundary of the overlapping region of the two input images, which often leads to the structural distortion of the stitched image.

3.4. Discussion

As mentioned above, even though there is no obvious structural distortion, Figure 1d still shows unnatural blending along the curb of the paved road resulted from inconsistent image brightness on both sides of the seam. Brightness/color correction can reduce such phenomenon to a certain extent. In the following experiments, the method in [29] is used to correct the brightness of the two original images shown in Figure 9a, and then the corrected images shown in Figure 9b are pre-aligned with the alignment approach used in the SPW method [13]. After that, the perceptual seam-cutting method [18] and the traditional seam-driven method based on color and texture information and the proposed method are used to stitch the two corrected images. The experimental results are given in Figure 9c–e.



Figure 8. The image stitching results of different methods and the used seams. (a) The perceptual seamcutting method in [18]. (b) The proposed method. (c) The final stitched image of SEAGULL dataset [17].

It can be found that there are obvious blur regions at the tree top in the red boxes, derived from the inappropriate position of the used seam in the process of Poisson fusion. Compared with the other two methods, the proposed method produces much less blurring. For the regions in the blue boxes, there exist obvious structural distortion at the bottom of the wall, as shown in partial enlargements of Figure 9c,d. The proposed method also produces some structural distortion in the region enclosed by the blue box, but the distortion is not obvious in subjective perception because the background here is thick leaves and the distance is relatively far. As shown in the green boxes, although the regions of paved road generated by the three methods are slightly different on both sides of the seam, the brightness difference is relatively small compared with that without brightness correction, so the naturalness of the paved road is improved. This indicates that brightness/color

correction helps to hide the seam and improve the quality of the stitched image when the brightness/color of the images to be stitched is inconsistent.

(b)



(c)

(a)

(**d**)

(e)

Figure 9. Image stitching results after brightness correction. (**a**) The original images to be stitched. (**b**) Images after brightness correction with the method in [29]. (**c**) The perceptual seam-cutting method [18]. (**d**) The traditional seam-driven method based on color and texture information. (**e**) The proposed method.

4. Conclusions

For images with large parallax, image misalignment and ghosting are the most challenging problems in their image stitching. In this paper, an optimized seam-driven image stitching method that additionally integrates scene depth information is proposed. Firstly, the input images to be stitched are aligned by using the single-perspective warps method so as to reduce the projection distortion on the basis of ensuring the accuracy of image alignment. Then, an energy function that integrates the depth information of the scene with the color and texture differences is defined so as to make the seam pass through the high similarity regions and bypass the prominent objects as much as possible. Based on the improved energy function, the graph cut optimization algorithm is used to find the seam. Finally, the Poisson fusion strategy is used to fuse the images and hide the seam. Experimental results have shown that the proposed method has a certain ability to understand the scene and finally generate more natural image stitching results, benefitting from the defined energy function which integrates the color, texture, and depth information of the scene. Moreover, when the brightness/color of the input images are inconsistent, the brightness/color correction of the images to be stitched can not only enhance the invisibility of the seam, but also improve the alignment accuracy. However, what kind of image features should be introduced into the energy function and how to effectively weighting the effect of different kinds of image features in the energy function is needed to be further studied, so as to improve the robustness of the proposed method, such as the cases of large depth of field scenes and small depth of field scenes. In the future, deep-learning-based image stitching will also be studied to further improve the generalization of the seam-driven image stitching method.

Author Contributions: Conceptualization, X.C., M.Y. and Y.S.; Methodology, X.C., M.Y. and Y.S.; Software, X.C.; Writing—original draft, X.C., M.Y. and Y.S.; Writing—review and editing, M.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 61871247, and in part by the Natural Science Foundation of Zhejiang Province under Grant LY21F010003. It was also sponsored by the K. C. Wong Magna Fund of Ningbo University.

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

References

- 1. Zhang, Y.; Lai, Y.-K.; Zhang, F.-L. Content-preserving image stitching with piecewise rectangular boundary constraints. *IEEE Trans. Visual. Comput. Graph.* **2021**, *27*, 3198–3212. [CrossRef] [PubMed]
- Aguiar, M.J.R.; Alves, T.d.R.; Honório, L.M.; Junior, I.C.S.; Vidal, V.F. Performance Evaluation of Bundle Adjustment with Population Based Optimization Algorithms Applied to Panoramic Image Stitching. *Sensors* 2021, 21, 5054. [CrossRef] [PubMed]
- 3. Cui, J.; Liu, M.; Zhang, Z.; Yang, S.; Ning, J. Robust UAV thermal infrared remote sensing images stitching via overlap-prior-based global similarity prior model. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* **2021**, *14*, 270–282. [CrossRef]
- 4. Liu, J.; Li, X.; Shen, S.; Jiang, X.; Chen, W.; Li, Z. Research on panoramic stitching algorithm of lateral cranial sequence images in dental multifunctional cone beam computed tomography. *Sensors* **2021**, *21*, 2200. [CrossRef] [PubMed]
- Guy, S.; Haberbusch, J.-L.; Promayon, E.; Mancini, S.; Voros, S. Qualitative comparison of image stitching algorithms for multi-camera systems in laparoscopy. J. Imaging 2022, 8, 52. [CrossRef]
- 6. Muñoz, L.; Díaz, C.; Orduna, M.; Ronda, J.I.; Pérez, P.; Benito, I.; García, N. Methodology for fine-grained monitoring of the quality perceived by users on 360VR contents. *Digit. Signal Process.* **2020**, *100*, 10–27. [CrossRef]
- Hosseinzadeh, S.; Jackson, W.; Zhang, D.; Mcdonald, L.; Macleod, C. A novel centralization method for pipe image stitching. *IEEE Sens. J.* 2021, 21, 11889–11898. [CrossRef]
- 8. Brown, M.; Lowe, D.G. Automatic panoramic image stitching using invariant features. *Int. J. Comput. Vis.* 2007, 74, 59–73. [CrossRef]
- Gao, J.; Kim, S.J.; Brown, M.S. Constructing Image Panoramas Using Dual-Homography Warping. In Proceedings of the 2015 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Colorado, CO, USA, 20–25 June 2011; pp. 49–56.
- 10. Zaragoza, J.; Chin, T.; Tran, Q.; Brown, M.S.; Suter, D. As-projective-as-possible image stitching with moving DLT. *IEEE Trans. Patern Anal. Mach. Intell.* **2014**, *36*, 1285–1298.
- Lin, C.C.; Pankanti, S.U.; Ramamurthy, K.N.; Aravkin, A.Y. Adaptive As-Natural-As-Possible Image Stitching. In Proceedings of the 2015 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1155–1163.
- Li, J.; Wang, Z.; Lai, S.; Zhai, Y.; Zhang, M. Parallax-tolerant image stitching based on robust elastic warping. *IEEE Trans. Multimed.* 2018, 20, 1672–1687. [CrossRef]
- Liao, T.; Li, N. Single-perspective warps in natural image stitching. *IEEE Trans. Image Process.* 2020, 29, 724–735. [CrossRef] [PubMed]
- 14. Shi, Z.; Wang, P.; Cao, Q.; Ding, C.; Luo, T. Misalignment-eliminated warping image stitching method with grid-based motion statistics matching. *Multimed. Tools Appl.* **2022**, *81*, 10723–10742. [CrossRef]
- 15. Gao, J.; Li, Y.; Chin, T.J.; Brown, M.S. Seam-Driven Image Stitching. In Proceedings of the 2013 Eurographics, Girona, Spain, 6–10 May 2013; pp. 45–48.
- 16. Huang, C.; Lin, S.; Chen, J. Efficient image stitching of continuous image sequence with image and seam selections. *IEEE Sens. J.* **2015**, *15*, 5910–5918. [CrossRef]

- Lin, K.; Jiang, N.L.F.; Cheong, M.D.; Lu, J. Seagull: Seam-Guided Local Alignment for Parallax-Tolerant Image Stitching. In Proceedings of the 2016 European Conference on Computer Vision—(ECCV), Amsterdam, The Netherlands, 8–16 October 2016; pp. 370–385.
- 18. Li, N.; Liao, T.; Wang, C. Perception-based seam cutting for image stitching. *Signal Image Video Process.* **2018**, *12*, 967–974. [CrossRef]
- Herrmann, C.; Wang, C.; Bowen, R.S.; Keyder, E.; Zabih, R. Object-Centered Image Stitching. In Proceedings of the 2018 European Conference on Computer Vision—(ECCV), Munich, Germany, 10–13 September 2018; pp. 846–861.
- 20. Wang, Z.; Yang, Z. Seam elimination based on Curvelet for image stitching. Soft Comput. 2019, 23, 5065–5080. [CrossRef]
- Agarwala, A.; Dontcheva, M.; Agrawala, M.; Drucker, S.; Colburn, A.; Curless, B.; Salesin, D.; Cohen, M. Interactive digital photomontage. ACM Trans. Graph. 2004, 23, 294–302. [CrossRef]
- 22. Boykov, Y.; Kolmogorov, V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. Patern Anal. Mach. Intell.* 2004, 26, 1124–1137. [CrossRef]
- Jung, K.; Hong, J. Quantitative assessment method of image stitching performance based on estimation of planar parallax. *IEEE Access* 2021, 9, 6152–6163. [CrossRef]
- Li, L.; Yao, J.; Xie, R.; Xia, M.; Xiang, B. Superpixel-Based Optimal Seamline Detection Via Graph Cuts for Panoramic Images. In Proceedings of the 2016 IEEE International Conference on Information & Automation, Ningbo, China, 1–3 August 2016; pp. 1484–1489.
- Seyed Mahdi Hosseini, M.; Sebastian, D.; Long, M.; Sylvain, P.; Yagız, A. Boosting Monocular Depth Estimation Models to High-Resolution Via Content-Adaptive Multi-Resolution Merging. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, WA, USA, 19–25 June 2021; pp. 9685–9694.
- Xu, J.; Hou, Y.; Ren, D.; Liu, L.; Zhu, F.; Yu, M.; Wang, H.; Shao, L. STAR: A Structure and texture aware retinex model. *IEEE Trans. Image Process.* 2020, 29, 5022–5037. [CrossRef]
- 27. Pérez, P.; Michel, G.; Andrew, B. Poisson image editing. ACM Trans. Graph. 2003, 22, 313–318. [CrossRef]
- Zhang, F.; Liu, F. Parallax-Tolerant Image Stitching. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 3262–3269.
- 29. HaCohen, Y.; Shechtman, E.; Goldman, D.B.; Lischinski, D. Non-rigid dense correspondence with applications for image enhancement. *ACM Trans. Graph.* 2011, *30*, 70. [CrossRef]