

## Article

# Quality-Aware Seam-Cutting and Content-Preserving Warping for Parallax Image Stitching

Jinliang Wang \*, Wenjuan Wang, Guobao Chen and Siwei Zhi

School of Information Engineering, Jiangxi Institute of Technology, Nanchang 330098, China; jinlaingwang07141996@126.com (W.W.); zsw\_xx@163.com (G.C.); siweizhi0627@126.com (S.Z.)

\* Correspondence: wanglang07141996@163.com

## Abstract

Parallax images often suffer from insufficient and unstable point features, making accurate alignment in overlapping regions extremely challenging. To address this issue, this paper proposes a robust image stitching framework based on content-preserving warping and quality-aware seam-cutting. The framework first estimates a global homography by combining point and line features to achieve reliable initial alignment. Subsequently, an energy function composed of alignment, distortion, and saliency terms is optimized to reduce geometric deformation while preserving image structures. After obtaining an initial seam using conventional seam-cutting, a structural similarity (SSIM)-based quality evaluation mechanism is introduced to identify misaligned regions and extract corresponding patches. To refine local alignment, the scale-invariant feature transform flow (SIFT flow) is employed to estimate dense correspondences, and a sigmoid-based smoothing constraint is adopted to avoid boundary artifacts. Finally, the optimized patches are integrated into the final result using a revised seam-cutting strategy with enhanced data penalty terms. Experimental results demonstrate that the proposed method achieves competitive performance, with favorable artifact suppression and structural preservation compared with mainstream stitching methods.

**Keywords:** image stitching; content-preserving warping; seam-cutting; SIFT flow

## 1. Introduction

Image stitching aims to merge multiple overlapping images into a single high-resolution panoramic image with a wide field of view. In image stitching, parallax refers to the apparent displacement of objects caused by camera translation and scene depth variation, which leads to inconsistent correspondences in overlapping regions. The homography is a  $3 \times 3$  projective transformation that maps points from one image plane to another and is globally valid only for planar scenes or pure camera rotation. Quasi-homography warping relaxes the global homography constraint by combining a global transformation with spatially varying local adjustments, thereby handling parallax more effectively while preserving the overall image structure. This technique has been widely applied in virtual reality, medical imaging, autonomous driving, UAV surveying, and other fields [1,2]. However, parallax caused by camera translation and scene depth variation remains a core challenge in practical applications. Traditional methods based on global homography often fail to achieve satisfactory alignment under large parallax, leading to ghosting, blurring, and structural distortion.



Academic Editor: Paolo Spagnolo

Received: 14 September 2026

Revised: 22 October 2026

Accepted: 23 October 2026

Published: 27 October 2026

**Copyright:** © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

At present, research on addressing parallax issues mainly focuses on two main directions: spatially varying warping [3] and seam-cutting methods [4]. Spatially varying warping subdivides the target image into multiple subregions and formulates the alignment as an energy minimization problem. However, since a homography is estimated for each subregion, this method may lead to homography inconsistency, causing lines in the final stitched image to appear curved. To preserve the overall structure of the image as much as possible, most studies tend to calculate the global homography between images. The global homography can effectively maintain the linear structure of the image but introduces projective distortion during the warping process. To alleviate these issues, researchers typically optimize the global homography to achieve better stitching results. In recent years, seam-cutting and seam-driven methods have attracted extensive attention due to their effectiveness in handling parallax problems [5,6]. These methods strive to precisely select the optimal stitching seam from a limited set of candidate seams in the overlapping regions. However, the efficacy of seam-cutting relies on accurate registration to search for reliable seams within the overlapping regions. Conventional seam-cutting methods may result in object cropping or duplication [7]. Existing deformation-based approaches suffer from homography inconsistency and structural distortion under large parallax, while most seam-cutting algorithms only focus on local seam optimization and ignore the global geometric constraints of images. Furthermore, mainstream stitching algorithms are mostly designed for a single type of parallax scenario, lacking adaptability to challenging scenes such as low-texture regions and complex foregrounds.

To address the aforementioned limitations, this paper proposes a novel framework for parallax image stitching. Leveraging established techniques as its foundation, the framework incorporates targeted enhancements, adaptive optimizations, and a synergistic integration of multiple modules. Each component is specifically redesigned and jointly optimized for large-parallax scenarios, allowing the framework to effectively address the fundamental challenges of parallax image stitching, including feature instability, geometric distortion, and stitching artifacts. The main contributions of this work are summarized as follows:

- (1) A dual-feature alignment strategy fusing SIFT point and LSD line features is developed to improve registration robustness in low-texture parallax scenes.
- (2) A unified multi-term energy function integrating alignment, distortion and saliency constraints is constructed. This function is adaptively optimized for parallax-induced geometric deformation, so as to suppress projection and perspective distortion while fully preserving image structures.
- (3) An SSIM-based seam quality evaluation mechanism combined with Otsu adaptive thresholding is designed to automatically localize misaligned patches in overlapping parallax regions.
- (4) A patch integration scheme coupling SIFT flow dense local alignment and sigmoid smoothing constraints is proposed to eliminate boundary artifacts while maintaining global structural consistency.

The rest of this paper is organized as follows: Section 2 reviews related work in the field of image stitching, with a focus on content-preserving warping, seam-cutting, and deep image stitching methods. Section 3 elaborates on the proposed image stitching method, which integrates content-preserving warping and stable seam-cutting to achieve visually natural and harmonious stitching for large-parallax images. Section 4 presents comprehensive quantitative, qualitative, and ablation experiments to fully validate the effectiveness of the proposed method. Finally, Section 5 concludes the paper and summarizes the main contributions.

## 2. Related Works

Conventionally, image stitching mainly involves three steps: image alignment, composition, and blending [8,9]. This section conducts an exhaustive literature survey centering on two core technical branches closely related to our work: content-preserving mesh warping and quality-aware seam-cutting, with a supplementary in-depth review of recent deep learning-based stitching paradigms.

### 2.1. Content-Preserving Warping

Parallax distortion remains a critical challenge in panoramic image stitching. Early content-preserving warping methods divided the scene into two regions and computed separate homography matrices, which were linearly combined to reduce parallax [10]. Li et al. further introduced a smoothly varying affine stitching domain and adopted multiple local homographies to handle complex parallax [11]. However, this may cause homography inconsistency and object distortion. To preserve image structures, the as-projective-as-possible (APAP) method uses a local parametric warping model and optimizes grid-wise homography via Gaussian weights [12]. Du et al. also explored similarity transformation for non-overlapping regions and derived global homography to suppress distortion [13]. Liao et al. proposed two single-perspective warping schemes: parametric warping and mesh-based warping, which optimize a unified energy function to suppress distortion and produce natural results [14]. Although these methods alleviate projective distortion, they rely heavily on accurate alignment of overlapping regions. A mixed behavior agent critic architecture was thus proposed to estimate robust seams and reduce parallax, using both point and line features with line constraints to maintain structure [15]. In follow-up improvements targeting low-texture scenes, Xiang et al. [16] proposed line-guided local warping with global similarity priors, embedding LSD line constraints to stabilize collinear structures. However, its performance degrades significantly when both point and line features are extremely scarce. Nonetheless, the above methods still depend on well-aligned overlaps. In contrast, seam-cutting compositing techniques have been independently developed to eliminate local parallax artifacts within overlapping regions, which will be systematically reviewed in the following subsection [17,18].

### 2.2. Seam-Cutting

Seam-cutting formulates optimal partition boundary extraction as discrete graph-cut energy minimization problems, with standard objective functions combining data fidelity, chromatic consistency, and spatial smoothness regularization terms [19]. Early pixel-level seam evaluation only calculated the single-channel intensity difference between corresponding pixels, easily misjudging visually smooth boundaries with minor pixel deviation. Li et al. [20] proposed perception-driven seam-cutting by introducing human visual weighting coefficients into the graph-cut loss, yet the optimization scope is limited within overlapping regions without global structural constraints. Zhang et al. incorporated visual perception into an energy minimization framework, applying S-shaped metrics to simulate color perception and saliency weights to locate visual focus [21]. Although this method yields perceptually favorable seams, it may not achieve global optimality. To address this issue, Jia et al. introduced projective invariants and feature correspondences to identify well-matched sub-regions for robust homography estimation, and embedded global collinear constraints into the objective function to reduce distortion and preserve global structure [22]. Recently, Chen et al. proposed an optimized seam-driven stitching strategy by fusing multi-dimensional features including depth, color and texture. For scenes with complex textures and continuous depth variations, they improved the energy function with depth constraints to guide seams away from salient objects, thus effectively improving

the stitching quality [23]. To overcome the limitation of pixel-only evaluation, Liao et al. [24] proposed quality evaluation-based iterative seam estimation for image stitching, extracting fixed-size local windows along candidate seams for secondary matching to locate misaligned zones. Liao et al. [25] further designed iterative patch quality evaluation for multi-view stitching, and Duan et al. [26] developed attentive deep image quality assessment for omnidirectional stitching. However, three common limitations exist in the patch-based seam-cutting literature: first, most existing schemes adopt manually fixed patch dimensions and static segmentation thresholds, lacking adaptive anomaly localization algorithms such as Otsu's inter-class variance maximization; second, patch assessment is treated as an isolated preprocessing step decoupled from alignment deformation, without bidirectional feedback between seam quality and global warping parameters; third, most of these methods focus on seam pixels themselves without considering the structural correlation of the surrounding regions in parallax scenarios.

In summary, seam-cutting algorithms are limited to local pixel/patch optimization, without the joint global dual-feature alignment and multi-term distortion suppression preprocessing. When low-texture scenes induce large registration residuals, standalone seam refinement cannot fundamentally eliminate parallax artifacts, which necessitates the integration of dual-feature warping and adaptive closed-loop seam evaluation proposed in this work.

### 2.3. Deep Image Stitching

In recent years, the rapid progress of deep learning has propelled data-driven image stitching, neural warping, and learning-based alignment to the forefront of research in this field. In contrast to conventional geometric approaches, deep learning-based methods are capable of extracting representative features from extensive datasets and learning transformation models, exhibiting competitive performance and distinct advantages in complex parallax scenarios and demanding scenes. Dynamic Time Warping (DTW) has been widely employed for sequence alignment and image matching tasks. Its differentiable counterpart, Soft-DTW, facilitates end-to-end learning for image registration by relaxing the original hard alignment loss and enabling effective optimization of network parameters [27]. Subsequently, Uncertainty-DTW further incorporates confidence estimation into the alignment pipeline [28], which effectively suppresses mismatches induced by noise and parallax, thereby bolstering robustness in unstable environments. In the domain of neural warping, several studies have adopted encoder–decoder architectures to extract multi-scale features [29], while others have combined convolutional neural networks with Transformer structures to achieve semantic-aware warping that adaptively adjusts deformation according to image content [30]. Although these approaches outperform traditional mesh-based warping methods, they generally demand large quantities of annotated data and exhibit limited generalization capability when applied to unseen real-world scenes. More recently, Nie et al. [31] proposed a parallax-tolerant unsupervised deep image stitching framework that leverages a flexible warping module and a seam-driven composition network to handle large-parallax scenarios without requiring ground-truth labels. This method demonstrates competitive performance on challenging datasets, representing a promising direction for learning-based stitching.

## 3. Proposed Method

As illustrated in Figure 1, the proposed robust image stitching framework comprises two core modules: image alignment and image composition. In the alignment process, we fuse point features (scale-invariant feature transform, SIFT) with line features (line segment detector, LSD) to estimate a reliable global homography matrix that adapts to parallax

scenes. On this basis, a content-preserving warping scheme constrained by a multi-term energy function is adopted to balance alignment accuracy, geometric distortion, and visual saliency, thereby effectively preserving image structures. In the composition process, we break through the limitations of conventional seam-cutting by introducing an SSIM-based quality assessment mechanism to identify misaligned regions. SIFT flow local alignment and sigmoid smoothing constraints are coupled to suppress artifacts. Finally, an optimized seam-cutting strategy is employed to fuse the refined image patches into a seamless and structurally intact stitching result.

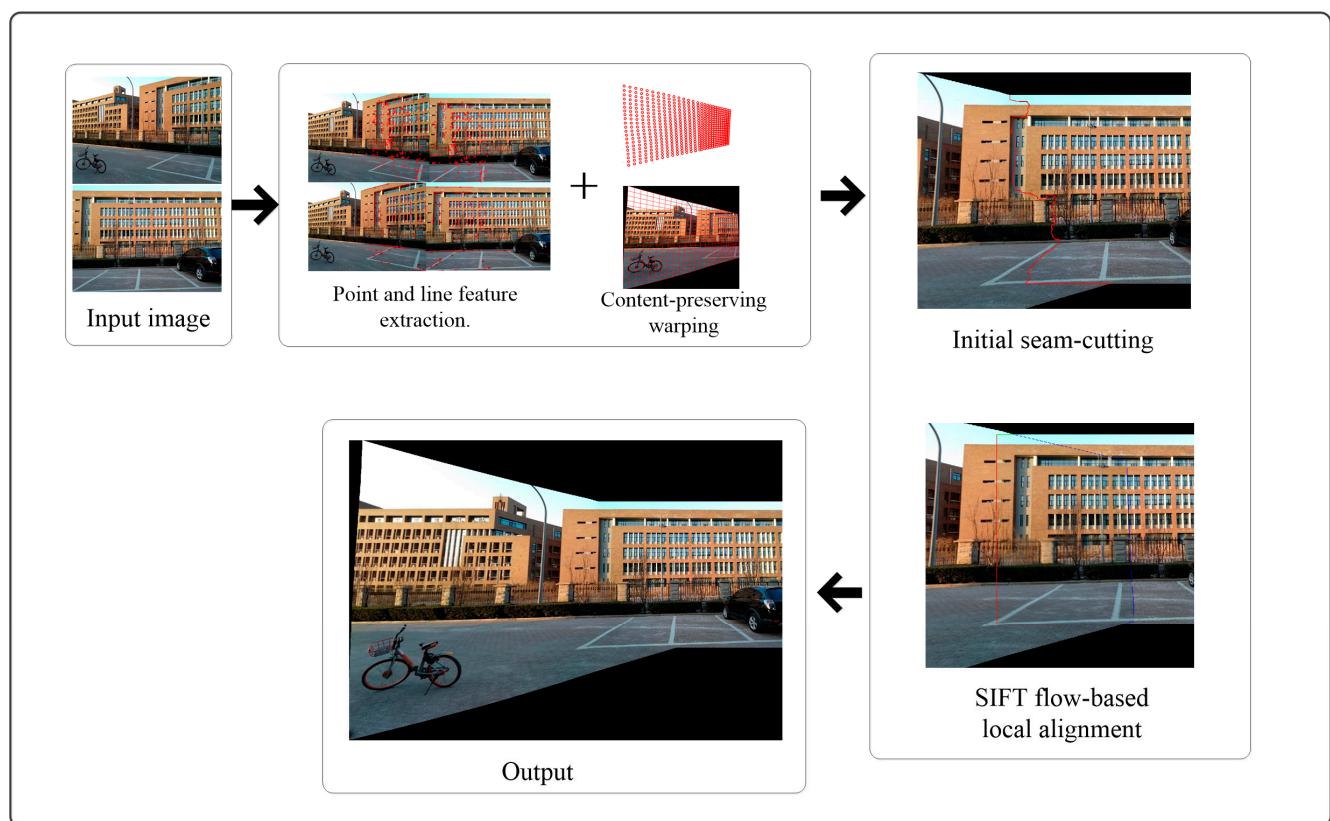


**Figure 1.** Overall flowchart of the proposed parallax image stitching framework, including dual-feature homography estimation, content-preserving warping, SSIM-based seam quality evaluation, SIFT flow local refinement and optimized seam integration.

### 3.1. Dual-Feature Global Homography Estimation

Let  $I$  and  $I'$  denote the target and reference images. Let  $p_i = [x_i, y_i, 1]$  and  $p'_i = [x'_i, y'_i, 1]$  be matched point pairs extracted by SIFT, and  $\{(l_j, l'_j)\}_{j=1,2,\dots,L}$  be line pairs extracted by LSD. The global dual-feature homography  $H_{dual}$  is estimated as:

$$H_{dual} = \underset{H}{\operatorname{argmin}} \sum_i \|p'_i - Hp_i\|^2 + \lambda \sum_j (\|a'_j - Ha_j\|^2 + \|b'_j - Hb_j\|^2), \quad (1)$$

where  $a_j$  and  $b_j$  are endpoints of line  $l_j$ , and  $\lambda$  is a balancing coefficient. The dual-feature fusion strategy is specifically optimized for parallax scenes with sparse point features, thereby enhancing the registration stability of the global homography model in complex parallax environments.

### 3.2. Content-Preserving Energy Optimization

**Point alignment term:** The point feature  $p$  can be represented as  $p = \omega V$  by combining vertex coordinates  $V = [V_1, V_2, V_3, V_4]^T$  of four mesh cells, where  $\omega = [\omega_1, \omega_2, \omega_3, \omega_4]$ .

After warping, the vertex coordinates become  $V = [\tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_4]^T$ , and the warped point  $\tilde{p}$  can be obtained by bilinear interpolation. Therefore, the point alignment term  $E_p$  is used to maximize the correspondence of point features in the input images. Its definition is as follows:

$$E_p = \sum_i \|\omega_i \tilde{V} - p'_i\|^2, \quad (2)$$

where  $p'_i$  denotes the matched point features.

**Line alignment term:** To further improve alignment, line correspondences are considered. Let  $l$  and  $l'$  denote the corresponding lines in the target and reference images. Let  $\{p_k\}$  denote the endpoints of all line segments, where  $p_k = \omega_{pk} V_{pk}$ ,  $V_{pk}$  denotes the vertex vector of mesh where  $p_k$  is located, and  $l'_j$  is denoted by the equation  $m_j x + n_j y + c_j = 0$ , where  $\sqrt{m_j^2 + n_j^2} = 1$  (Please refer to Reference [14]). The line alignment term  $E_l$  is defined as:

$$E_l = \sum_{j,k} \|(l'^T_j \cdot \omega_{pk} \tilde{V}_{pk}) / (\sqrt{m_j^2 + n_j^2})\|^2, \quad (3)$$

**Distortion term:** To suppress the distortion problem between images, previous methods have shown that quasi-homography warping alleviate it to a certain extent. The projective and perspective consistencies are maintained via optimizing the slope of the cross-line feature.  $l_u$  denotes a horizontal line that preserves its slope after warping, and  $l_v$  divides the overlapping and non-overlapping regions. The slopes of parallel lines  $l_u$  can be expressed as:

$$k_1 = -\frac{h_7}{h_8}, \quad s(x, y, k_1) = \frac{h_4 h_8 - h_5 h_7}{h_1 h_8 - h_2 h_7}, \quad (4)$$

where  $h_1, h_2, h_4, h_5, h_7$  and  $h_8$  denote the elements of the global homography matrix  $H$  in Equation (1).  $k_1$  and  $k_2$  denote the slopes of the horizontal and vertical lines, respectively, and  $l_v$  can be set to the dividing line according to slope  $k_1$  that is closest to the border between the overlapping region and the non-overlapping regions. Therefore, the slope of  $l_v$  can be represented as perpendicular to  $k_1$ :

$$k_1 \cdot k_2 = -1, \quad s(x, y, k_1) \cdot s(x, y, k_2) = -1, \quad (5)$$

The perpendicularity condition in Equation (5) is expressed as the product of the two slopes being  $-1$  ( $s(x, y, k_1) \cdot s(x, y, k_2) = -1$ ). The intersection line can be obtained through Equations (4) and (5).

In fact, each pixel can be represented by  $l_u$  and  $l_v$  in the target image and can be represented by  $l'_u$  and  $l'_v$  after warping. The cross-line feature can be represented by  $\{(l'_u, l'_u)\}$  and  $\{(l'_v, l'_v)\}$ . The perspective distortion term is defined as:

$$\begin{aligned} E_{ps} &= \sum_{i=1} \sum_{k=1} |\langle \omega_{p_{k+1}^{ui}} \tilde{V}_{p_{k+1}^{ui}} - \omega_{p_k^{ui}} \tilde{V}_{p_k^{ui}}, \vec{n}_i^u \rangle|^2 \\ &+ \sum_{j=1} \sum_{k=1} |\langle \omega_{p_{k+1}^{vj}} \tilde{V}_{p_{k+1}^{vj}} - \omega_{p_k^{vj}} \tilde{V}_{p_k^{vj}}, \vec{n}_j^v \rangle|^2 \\ &+ \sum_{j=1} \sum_{k=1} \|\omega_{p_k^{vj}} \tilde{V}_{p_k^{vj}} + \omega_{p_{k+2}^{vj}} \tilde{V}_{p_{k+2}^{vj}} - 2\omega_{p_{k+1}^{vj}} \tilde{V}_{p_{k+1}^{vj}}\|^2, \end{aligned} \quad (6)$$

where  $\vec{n}_i^u$  and  $\vec{n}_j^v$  are the normal vectors of  $l_u^i$  and  $l_v^j$ . As shown in Figure 2, the first row of (6) represents the slopes of the warped horizontal line. To maintain slope consistency, the inner product of the slope and the normal vector is constrained. Specifically, the first two parts keep the slopes of  $l_u^i$  and  $l_v^j$ , and the last preserves the ratios of lengths on  $l_v^i$  (see Figure 2). Taking the green line as an example, the  $(\omega_{p_{k+1}^{ui}} \tilde{V}_{p_{k+1}^{ui}} - \omega_{p_k^{ui}} \tilde{V}_{p_k^{ui}})$  in the first row of Equation (6) can be regarded as the slope of horizontal lines after warping, and we obtain the inner product of the yellow line after warping and the normal vector before transformation to ensure the slope [14].

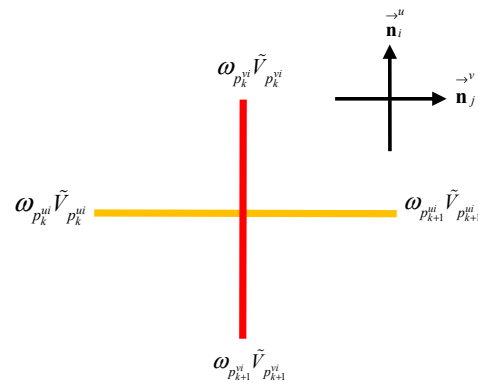


**Figure 2.** Illustration of cross-line features for projective and perspective distortion constraints. The red and yellow lines denote vertical and horizontal lines, respectively.

Projective distortion can be further suppressed as follows. The term in Equation (7) penalizes the second-order deviation of consecutive warped vertices along the vertical lines.

$$E_{pj} = \sum_{i=1} \sum_{k=1} \|\omega_{p_k^{uj}} \tilde{V}_{p_k^{uj}} + \omega_{p_{k+2}^{uj}} \tilde{V}_{p_{k+2}^{uj}} - 2\omega_{p_{k+1}^{uj}} \tilde{V}_{p_{k+1}^{uj}}\|^2. \quad (7)$$

Saliency term: To avoid line bending in non-overlapping regions, a reasonable saliency term is adopted. The salient lines are denoted as  $\{l_k^s\}_{k=1}$ , where  $l_k^s$  is a set of points. The saliency term is defined as:

$$E_s = \sum_{k=1} \sum_{j=1} \|\langle \omega_{p_{j+1}^k} \tilde{V}_{p_{j+1}^k} - \omega_{p_j^k} \tilde{V}_{p_j^k}, \vec{n}_k \rangle\|^2, \quad (8)$$

where  $\vec{n}_k$  is the norm of  $l_k^s$ . The constructed energy function integrates point alignment, line alignment, distortion suppression, and visual saliency constraints, and is adaptively optimized for geometric distortion caused by parallax.

The above energy terms are integrated into a unified energy optimization problem:

$$E_{energy} = E_p + \lambda_l E_l + \lambda_{ps} E_{ps} + \lambda_{pj} E_{pj} + \lambda_s E_s, \quad (9)$$

where  $\lambda_l, \lambda_{ps}, \lambda_{pj}$  and  $\lambda_s$  are weighting coefficients for each energy term.

### 3.3. Initial Seam-Cutting

Let  $I_0$  and  $I_1$  represent the target and reference images, respectively. Let  $\mathbf{L} = \{0, 1\}$  be a label set, where “0” corresponds to  $I_0$  and “1” corresponds to  $I_1$ . The optimal stitching seam is formulated as a label energy function optimization problem, in which a label  $l_p \in \mathbf{L}$  is assigned to each pixel of  $p \in \mathbf{P}$ . The label energy function is defined as:

$$E(l) = \sum_{p \in \mathbf{P}} D_p(l_p) + \sum_{(p,q) \in \mathbf{N}} S_{p,q}(l_p, l_q), \quad (10)$$

where  $\mathbf{N} \subset \mathbf{P} \times \mathbf{P}$  is the set of neighboring pixel pairs and  $\mathbf{P}$  denotes the overlapping region.  $D_p$  is the data term, which represents the cost of assigning labels, and  $S_{p,q}$  is the smoothness term in conventional seam-cutting.

### 3.4. SSIM-Based Seam Quality Evaluation

After aligning the target and reference images, we first compute an initial seam  $S$  using the seam-cutting method. The Structural Similarity Index (SSIM) [32] is then used to evaluate each pixel along the seam. For each pixel  $p_i$  on the seam, a local patch centered at  $p_i$  is extracted, and the SSIM value between corresponding patches of the input images is calculated. The evaluation error for pixel  $p_i$  is defined as:

$$Q(p_i) = E\left(\frac{1 - \text{SSIM}(I_0(p_i), I_1(p_i))}{2}\right), \quad (11)$$

where the SSIM index ranges from  $[-1, 1]$  and equals 1 only when two local patches are identical. Here,  $E(\cdot)$  denotes the mean operator over the local patch centered at  $p_i$ . Figure 3 shows the stitching seam results of the traditional method and the proposed method. As can be seen from the Y-axis value distribution in Figure 4, our method has a lower cost for pixels along the seam, thereby generating a more stable seam. Traditional seam evaluation methods only compute pixel-level color or grayscale costs along the seam, whereas our method employs local image patches as the evaluation carrier and utilizes SSIM to characterize structural differences. X-axis: pixel index along the stitching seam; Y-axis: local patch-based SSIM cost.



**Figure 3.** Comparison of the optimal stitching seams. (a) Seam generated by conventional seam-cutting; (b) Seam generated by the proposed quality-aware seam-cutting.

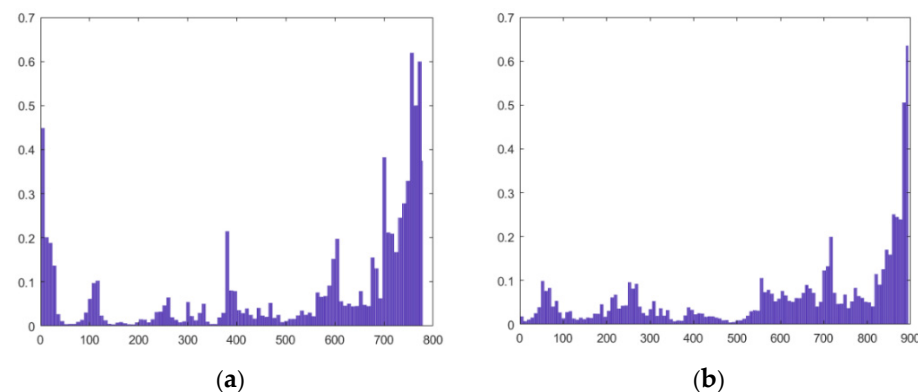


**Figure 4.** Local patch-based SSIM quality evaluation along the stitching seam. (a) Conventional seam-cutting; (b) The proposed quality-aware seam. The horizontal axis is the pixel index along the seam, and the vertical axis denotes the local patch SSIM cost.

Although existing seam evaluation and local block optimization methods also adopt block-based evaluation strategies, these works are fundamentally different from our approach [24–26]. First, most patch-based seam evaluation methods merely compute independent structural scores for local regions and treat them as static judgment criteria, without incorporating adaptive thresholding algorithms to achieve automatic localization of anomalous regions. In this paper, we integrate Otsu's thresholding method with SSIM evaluation results to realize intelligent localization of problematic regions in parallax stitching. Second, previous seam quality assessments served as a preprocessing step, whereas our SSIM evaluation results directly guide the SIFT-flow local alignment and sigmoid smoothing modules, forming a closed-loop pipeline from seam quality detection to artifact removal. Third, traditional patch evaluations only focus on the seam pixels themselves, while our

scheme considers the structural correlation of surrounding regions in parallax scenarios, which better conforms to the visual characteristics of parallax images.

### 3.5. Seam Integration

Since line features are introduced as alignment constraints, the initial seam usually contains only partial misalignment even for large-parallax images. Our goal is to correct these misaligned regions. A reasonable seam should have relatively small error  $Q$  for all pixels. Therefore, pixels with large errors are extracted from the seam for further alignment refinement. Specifically, if  $\max(Q) \leq k\text{mean}(Q)$  is satisfied, where  $k = 1.5$  is a constant, the seam is regarded as reasonable, and  $Q$  denotes the set of evaluation errors along the seam. Otherwise, a threshold  $\tau$  is calculated via Otsu's method [33], and pixels with  $p_i$  are marked as misaligned.

The separated patches are pre-aligned, so the parallax between these patches can be ignored. SIFT flow is then used to establish dense correspondences between pixels in the patches [34]. Theoretically, for each pixel  $p_j$  in  $I_0^i$ , there exists a shift vector  $V^i(p_j)$  such that:

$$\|I_0^i(p_j + V^i(p_j)) - I_1^i(p_j)\| = 0. \quad (12)$$

The two patches are then aligned using vector  $V^i$ . However, if the boundary pixels of the patch have large shift vectors, merging the aligned patches into the pre-aligned result will introduce severe boundary artifacts. Therefore, the vector flow  $V^i(p_j)$  is smoothly modified to  $f(V^i(p_j))$  through a sigmoid function, as shown in Figure 5. The smoothing function is defined as:

$$f(t) = \frac{1}{1 + e^{-\alpha(t-0.5)}}, \quad (13)$$

where  $t$  is the normalized x-coordinate of the pixel in the patch, which is normalized to  $[0, 1]$  and  $\alpha$  controls the steepness of the sigmoid curve of  $f$ .

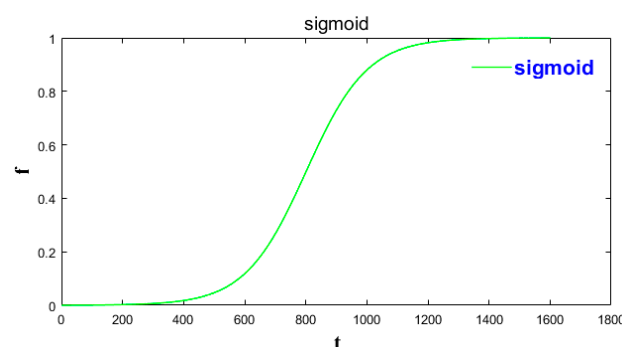


**Figure 5.** Curve of the sigmoid smoothing function applied to SIFT flow displacement vectors for boundary artifact suppression.

The separated patches are only partially aligned by the above method and still need to be composited via seam-cutting. Therefore, the data penalty term is modified as:

$$\begin{cases} D_{pj}(0) = 0, D_{pj}(1) = \infty, & \text{if } p \in \partial I_0 \cap \partial \mathbf{P} \\ D_{pj}(0) = 0, D_{pj}(1) = \infty, & \text{if } p \in \partial I_1 \cap \partial \mathbf{P} \\ D_{pj}(0) = \infty, D_{pj}(1) = 0, & \text{otherwise} \end{cases}, \quad (14)$$

where  $\infty$  is a very large penalty to avoid mislabeling and  $\partial I_k \cap \partial \mathbf{P}$  denotes the intersection of the boundary of the  $k$ -th patch and the boundary of the overlapping regions. In fact, the data term  $D_p(l_p)$  fixes the endpoints of the seam as the intersections of the two colored polylines. Since  $\partial$  is set to a sufficiently large value, its exact magnitude does not affect the optimal seam as long as it dominates the smoothness term; consequently, misaligned labels are always excluded from the solution.

The above seam integration process achieves the organic coupling of SSIM quality assessment, Otsu's threshold segmentation, SIFT flow local alignment, and sigmoid smoothing constraints. Each module mutually constrains and complements the others, forming a complete closed-loop optimization pipeline for parallax seam correction.

## 4. Experiments

To verify the effectiveness and advantages of the proposed method, comparative experiments are conducted against five representative stitching baselines, namely APAP [12], PBS [20], SPW [14], UDIS [31], and LAPM [25]. The experiments consist of two parts: quantitative evaluation on the public benchmark UDIS-D [35], and qualitative visual comparison on multi-scene image pairs. The UDIS-D benchmark contains 1106 test image pairs with a resolution of  $512 \times 512$ , which is divided into Easy, Moderate, and Hard subsets according to the degree of parallax. The qualitative test image pairs are collected from the public datasets of the aforementioned baseline methods, covering diverse real-world scenarios including indoor, outdoor, nighttime, low-light, snowy and zoom scenes.

To guarantee fair comparison, all input image pairs are uniformly resized to  $512 \times 512$  prior to stitching. SIFT [36] and LSD [37] are adopted for point and line feature extraction, respectively. All hyperparameters of the competing baselines are strictly configured in accordance with their original publications, and no extra manual tuning is performed. All algorithms are executed on the same hardware platform, a PC equipped with an Intel i7-9750H 2.6 GHz CPU, and all runtime results are averaged over three independent runs. For the proposed method, the patch size for SSIM evaluation is set to  $21 \times 21$ , and the constants  $k$  and  $\alpha$  are set to 1.5 and 8, respectively. Besides,  $\lambda_l, \lambda_{ps}, \lambda_{pj}$  and  $\lambda_s$  in Equation (9) are set to 5, 5, 50, 5. The rationale for these parameter settings is further analyzed in Section 4.5.

### 4.1. Quantitative Comparison

To evaluate the alignment performance of these warping methods, we conduct a quantitative comparison of all competing methods and the proposed on the standard public benchmark UDIS-D [36]. This benchmark contains 1106 test image pairs of  $512 \times 512$  resolution. Following the official evaluation protocol of UDIS-D, the test set is divided into Easy, Moderate, and Hard subsets according to the degree of parallax. All input image pairs are resized to  $512 \times 512$  before stitching. For fair comparison, point features are detected by SIFT [35] and line features by LSD, and all parameters of the compared methods are strictly set according to their original papers. The results are shown in Table 1. We adopt RMSE, PSNR, and SSIM [32] as image quality evaluation metrics to comprehensively compare alignment quality. Specifically, RMSE measures pixel-level alignment error, while PSNR and SSIM reflect pixel fidelity and structural similarity, respectively. Since UDIS-D does not provide ground-truth panoramas, these metrics are computed on the overlapping regions of the warped image pairs, using the warped reference image as the reference. For each image pair, we warp the reference and target images using the estimated transformation, and then compute RMSE, PSNR, and SSIM between the two warped images only within their overlapping region. In addition, we select specific regions from both the overlapping and non-overlapping areas to verify the effectiveness of the proposed method, as shown in Figure 6; RMSE, PSNR, and SSIM are also computed over the corresponding cropped regions with the warped reference image as the reference. The results for each subset (Easy, Moderate, and Hard) and the entire test set (Average) are averaged over all image pairs. For a few challenging image pairs where a method fails to produce any meaningful result, the failed cases are excluded from its average computation.

To further analyze the computational overhead, all algorithms are evaluated under identical hardware conditions. Table 2 lists the total runtime of competing stitching algorithms,

together with the module-wise time breakdown of the proposed method. As shown in Table 2, lightweight warping-based approaches such as APAP and PBS achieve fast inference. SPW and LAPM introduce extra mesh-optimization steps and incur higher computational costs. UDIS, as a deep-learning-based method, suffers from heavy CPU computation overhead in the absence of GPU acceleration. Within our framework, dual-feature homography estimation and multi-term energy optimization consume moderate computational resources. The major bottleneck originates from the SIFT-flow dense correspondence step for local patch alignment, which accounts for the majority of the total runtime. Although our method requires longer computation time than most baseline methods, it achieves superior stitching quality in challenging large-parallax, low-texture, and complex-foreground scenes. All timing results are averaged over three independent runs on the UDIS-D benchmark.

**Table 1.** Quantitative comparison of RMSE, PSNR and SSIM on the UDIS-D benchmark, divided into Easy, Moderate, Hard and Average subsets.

| Method | PSNR  |          |       |         | SSIM  |          |       |         | RMSE |          |      |         |
|--------|-------|----------|-------|---------|-------|----------|-------|---------|------|----------|------|---------|
|        | Easy  | Moderate | Hard  | Average | Easy  | Moderate | Hard  | Average | Easy | Moderate | Hard | Average |
| APAP   | 23.87 | 20.78    | 18.93 | 21.19   | 0.822 | 0.602    | 0.526 | 0.650   | 1.52 | 1.94     | 2.52 | 1.99    |
| PBS    | 24.87 | 21.95    | 19.43 | 22.08   | 0.848 | 0.551    | 0.453 | 0.617   | 1.92 | 2.39     | 2.43 | 2.25    |
| SPW    | 26.76 | 23.56    | 20.54 | 23.62   | 0.877 | 0.664    | 0.567 | 0.703   | 1.36 | 1.71     | 2.26 | 1.78    |
| UDIS   | 26.93 | 23.03    | 21.93 | 23.96   | 0.834 | 0.642    | 0.546 | 0.674   | 1.47 | 1.88     | 2.43 | 1.93    |
| LAPM   | 27.36 | 23.24    | 21.85 | 24.15   | 0.901 | 0.677    | 0.573 | 0.717   | 1.35 | 1.78     | 2.25 | 1.79    |
| Ours   | 29.69 | 25.17    | 23.78 | 26.21   | 0.921 | 0.744    | 0.648 | 0.771   | 1.19 | 1.51     | 2.78 | 1.83    |



**Figure 6.** Magnified local comparisons of representative image pairs, including the target image and the warping results of all competing methods and the proposed method.

**Table 2.** Average runtime comparison on the UDIS-D benchmark (seconds). Hardware: Intel i7-9750H 2.6 GHz CPU, without GPU acceleration. The right part lists the runtime breakdown of each module in the proposed framework.

| Baseline Algorithms | Runtime (s) | Module Decomposition                | Runtime (s) |
|---------------------|-------------|-------------------------------------|-------------|
| APAP                | 5.26        | Dual-feature global homography      | 4.14        |
| PBS                 | 6.71        | Multi-term energy optimization      | 6.07        |
| SPW                 | 7.34        | SSIM evaluation + Otsu thresholding | 3.86        |
| LAPM                | 10.12       | SIFT-flow dense patch alignment     | 7.22        |
| UDIS                | 12.47       | Sigmoid smoothing + seam fusion     | 2.54        |
|                     |             | Ours (Total)                        | 23.83       |

#### 4.2. Qualitative Results

In this section, comparative experiments are carried out on ten groups of multi-scene images collected from public datasets. Five mainstream stitching algorithms, namely APAP, PBS, SPW, UDIS, and LAPM, together with the proposed method, are selected for visual

comparison. Three representative test cases corresponding to Figure 7, Figure 8, and Figure 9 (campus, school, and building scenes) are analyzed in detail, covering small-parallax scenes, scenes with complex foreground, and large-parallax scenes with low-texture regions, respectively.

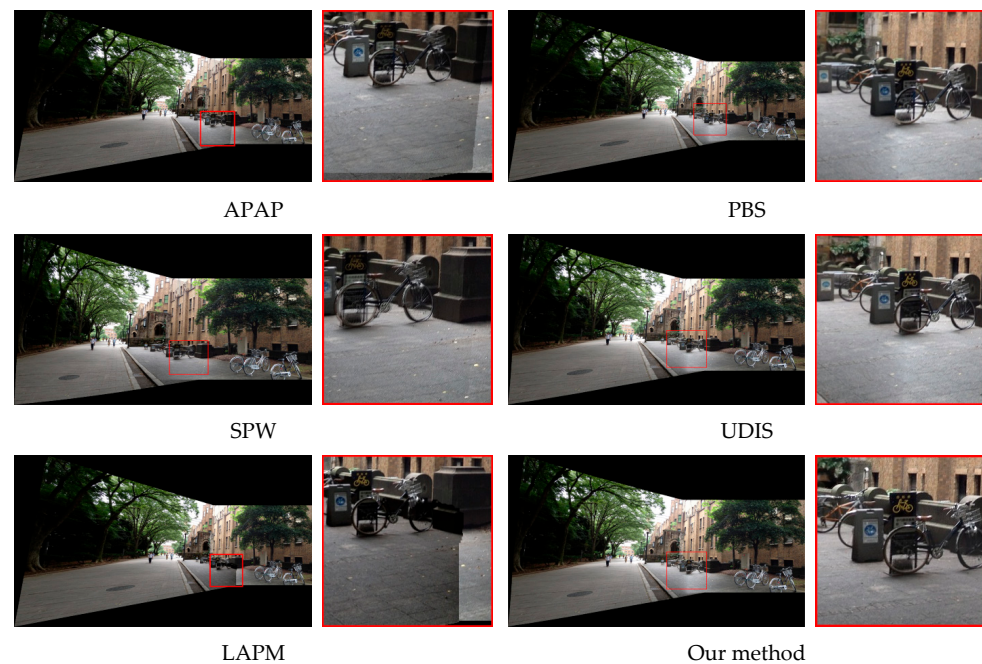


**Figure 7.** Compare the experimental results with other methods for the campus image pair.

Figure 7 presents the visual stitching results of APAP, PBS, SPW, UDIS, LAPM, and the proposed method for the campus image pair. For better comparison, partial details within the rectangular boxes are highlighted and magnified. Obvious bending distortion of object contours in the overlapping regions can be observed for both APAP and PBS, along with slight artifacts around the bicycle wheels. In addition, the LAPM algorithm shows distinct brightness inconsistency along the stitching seam. SPW yields relatively smooth seams but introduces local structural distortion in non-overlapping regions. Although UDIS can effectively suppress ghosting in overlapping areas, it causes noticeable color inconsistency. As can be seen from Figure 7, the proposed method can effectively eliminate artifacts. The stitching seam becomes nearly invisible with uniform brightness transition over the whole image, which generates stitching results with more natural visual appearance.

Figure 8 shows the experimental results of APAP, PBS, SPW, UDIS, LAPM, and the proposed method for the school image pair. For comparison purposes, partial details within the red rectangular boxes are magnified. Severe warping and structural distortion of wheel contours can be observed in the non-overlapping regions for APAP, PBS, and LAPM. The SPW algorithm achieves better global-structure preservation yet suffers from slight misalignment at the bicycle frame. UDIS mitigates distortion in non-overlapping zones at the cost of reduced clarity in texture-rich regions. Encouragingly, compared with the above competing algorithms, the proposed method can effectively preserve object integrity and yield bicycle geometric contours with more complete and natural shapes.

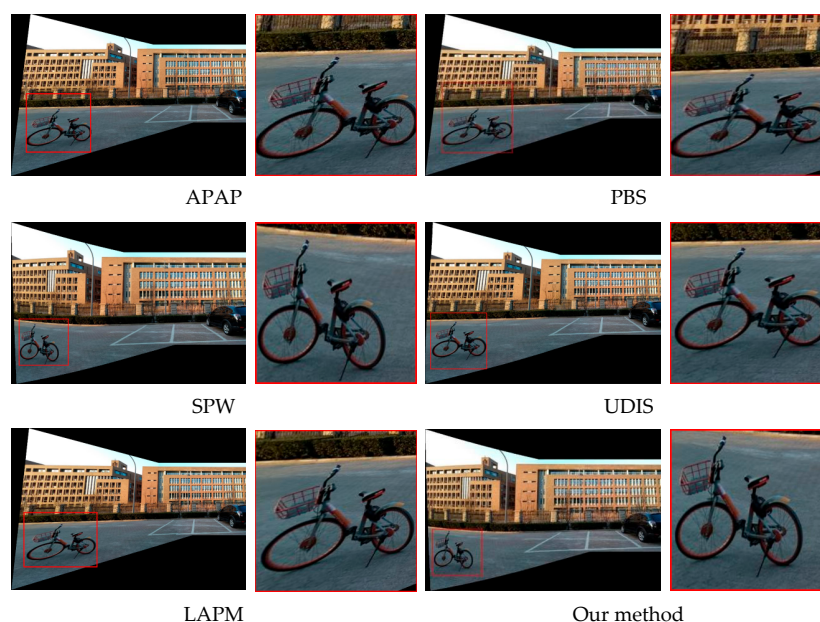


**Figure 8.** Comparison of stitching results for the school image pair.

Finally, Figure 9 presents the stitching results of six algorithms for the building image pair. The red and blue rectangular boxes show magnified local details of the non-overlapping and overlapping regions, respectively. Obvious artifacts and blurring can be observed in the magnified regions for APAP, which yields unsatisfactory stitching quality. PBS effectively removes artifacts in the overlapping zone yet introduces prominent deformation of windows in the non-overlapping region. SPW achieves competitive performance in preserving straight-line structures, but residual ghosting remains near object boundaries. UDIS obtains decent alignment, while slight blurring appears in high-frequency regions. Although LAPM mitigates distortion and misalignment to some extent, slight alignment inconsistencies still exist within the overlapping zone. Encouragingly, compared with the above algorithms, the stitching result of the proposed method can not only effectively eliminate artifacts in overlapping regions, but also suppress distortions in non-overlapping regions.

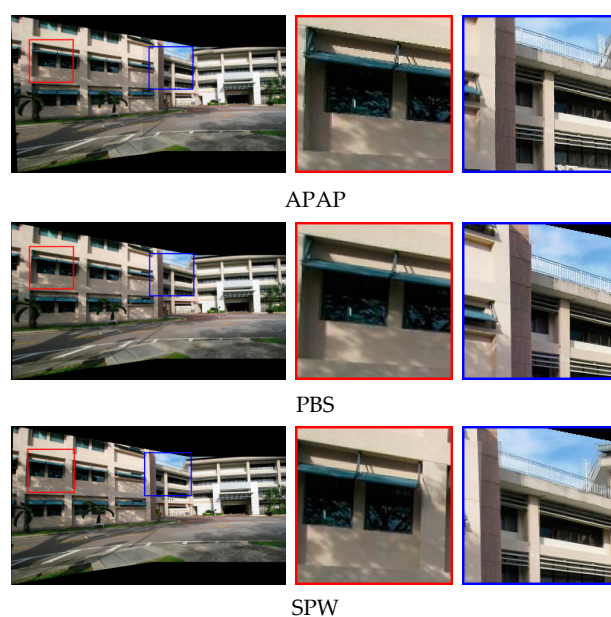


**Figure 9.** Cont.

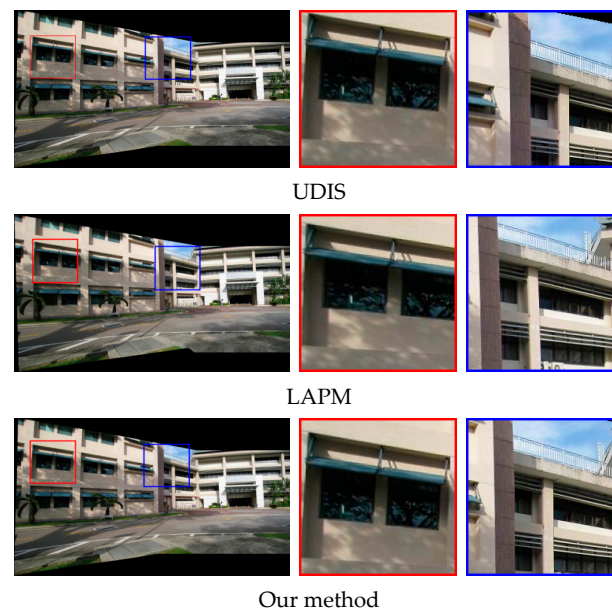


**Figure 9.** Comparison of stitching results for the building image pair.

#### 4.3. Ablation Experiments

To verify the effectiveness of each module in our method, we conduct ablation studies and present the averaged quantitative results on the overlapping regions of test samples in Table 3. To further validate the independent contribution of each core module in the proposed stitching pipeline, we conduct strict single-variable controlled ablation experiments on the test images, where only one module is incrementally added in each group while all other hyperparameters remain fixed to eliminate irrelevant interference. In addition, the corresponding visual comparison results for a representative scene are shown in Figure 10. The experimental results are arranged in four rows from top to bottom, with each row showing (from left to right) the final stitching result, the magnified local details of non-overlapping regions, and the magnified local details of overlapping regions. The first row adopts only the dual-feature global homography and the traditional seam-cutting, causing severe geometric distortion in the non-overlapping regions. In the second row, a multi-constraint energy function and SSIM-based quality assessment are further introduced. As a result, the structural distortion is alleviated to a certain extent; however, slight artifacts remain in the human figures within the overlapping region. In the third row, after incorporating Otsu's method and SIFT flow correction, the misalignment in the overlapping region is suppressed to some degree, yet the overall performance still falls short of the desired effect. Finally, with the addition of the Sigmoid smoothing module in the fourth row, the misalignment in the overlapping region is effectively eliminated, while the complete scene geometry in the non-overlapping region is well preserved. In summary, the progressively improved visual quality across the ablation groups substantiates the effectiveness of each core module within the proposed parallax-tolerant stitching framework.

**Table 3.** Quantitative ablation results of the proposed framework, averaged over all test image pairs, evaluated in the overlapping region using RMSE, PSNR and SSIM.

| Configuration                                                       | RMSE | PSNR  | SSIM  |
|---------------------------------------------------------------------|------|-------|-------|
| Baseline: Dual-feature homography + traditional seam-cutting        | 2.01 | 21.37 | 0.412 |
| Baseline + Multi-term energy optimization + SSIM quality evaluation | 1.82 | 22.65 | 0.438 |
| Baseline + Otsu thresholding + SIFT flow                            | 1.64 | 23.89 | 0.462 |
| Full Proposed Framework                                             | 1.49 | 25.17 | 0.744 |

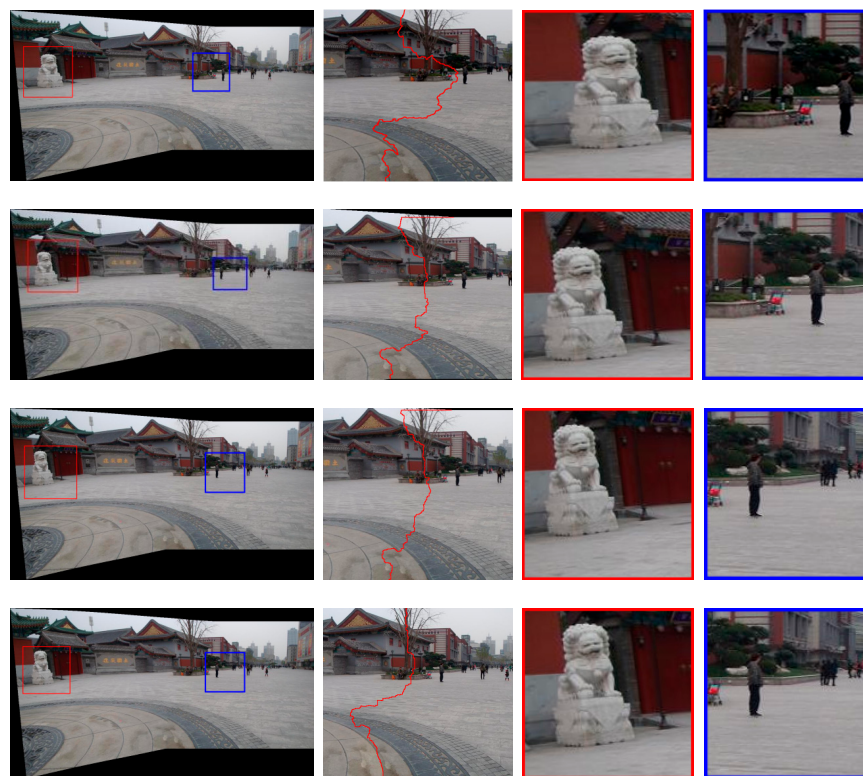


**Figure 10.** Ablation study on a representative scene. The results from top to bottom are the method adopting dual-feature global homography and traditional seam-cutting, the method introducing the multi-constraint energy function and SSIM quality evaluation, the method combining Otsu's method and SIFT flow, and our proposed method. From left to right are the final results, the magnified local details of non-overlapping regions, and the magnified local details of overlapping regions.

#### 4.4. Limitations and Failure Cases

Despite the promising performance demonstrated across various parallax scenarios, the proposed method still exhibits certain limitations that require further investigation. Firstly, in scenes with extremely severe parallax where the depth disparity exceeds the capacity of our dual-feature global homography initialization, the subsequent energy optimization may converge to suboptimal local minima, resulting in residual misalignment in the overlapping regions (Figure 11). Secondly, the presence of moving objects (e.g., pedestrians, vehicles) in the overlapping areas violates the static scene assumption underlying the homography-based alignment, which may introduce ghosting artifacts that cannot be fully eliminated by seam-cutting alone. Thirdly, in low-texture or repetitive-pattern scenes where both SIFT and LSD fail to extract sufficient reliable features, the initial global homography estimation becomes unstable, and the performance of our method degrades accordingly, as evidenced by the slightly higher RMSE on the low-texture “wall” and “corridor” datasets compared with the full-feature scenes. Fourthly, significant illumination variations between input images can impair the SSIM-based quality evaluation, as the structural similarity metric is inherently sensitive to luminance changes, potentially leading to suboptimal patch selection. Finally, the current framework requires reasonably well-aligned initial homography to ensure the convergence of SIFT flow local refinement; in cases where the initial alignment residual exceeds the typical displacement range of SIFT flow, the dense correspondence estimation may fail to capture the true misalignment. Addressing these limitations through robust feature augmentation, dynamic scene handling, and illumination-adaptive quality assessment constitutes an important direction for future work. In addition, the SIFT-flow dense matching module constitutes the main computational bottleneck of our pipeline,

leading to relatively high overall runtime. Future work will explore lightweight dense matching or down-sampling strategies to reduce computational cost, facilitating practical real-world deployment.

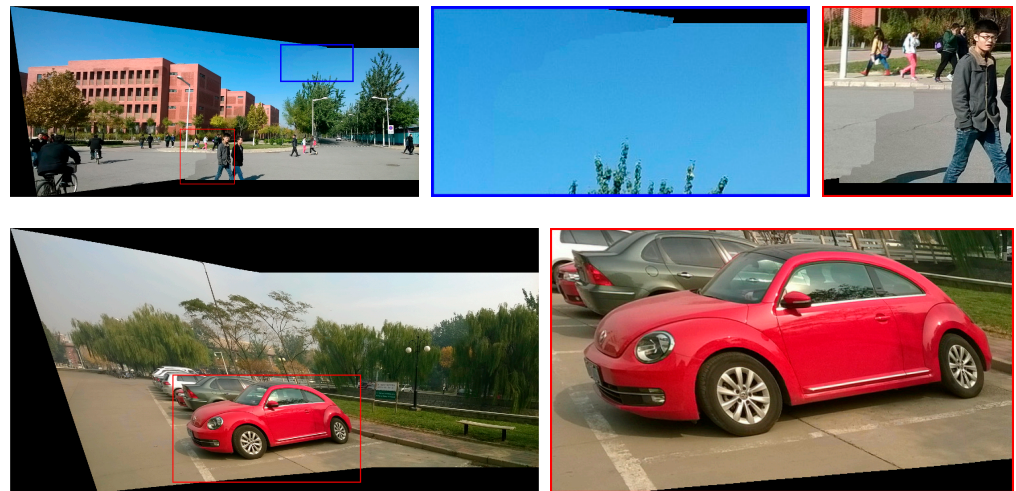


**Figure 11.** Failure cases of the proposed method. Green and red boxes mark residual misalignment artifacts caused by extreme parallax and moving objects.

#### 4.5. Parameter Sensitivity Analysis

To justify the parameter selection in Section 4.1, we further conduct a sensitivity analysis on the two most influential parameters of the proposed framework: the weight  $\lambda_{pj}$  of the projection-distortion term in Equation (9) and the patch size used in the SSIM-based quality evaluation. Figure 12 reports the average RMSE and SSIM over all test image pairs, where one parameter is varied over a wide range while the others are fixed at their default values ( $\lambda_l = 5$ ,  $\lambda_{ps} = 5$ ,  $\lambda_{pj} = 50$ ,  $\lambda_s = 5$ , patch size  $21 \times 21$ ,  $k = 1.5$ ,  $\alpha = 8$ ). It can be observed that the proposed method achieves the lowest RMSE and the highest SSIM at the default settings, and the performance remains stable when  $\lambda_{pj}$  varies from 10 to 70 and the patch size varies from  $9 \times 9$  to  $33 \times 33$ . These results confirm that the empirically determined parameter values are robust across different scene types and that the proposed method is not sensitive to the exact parameter choice.

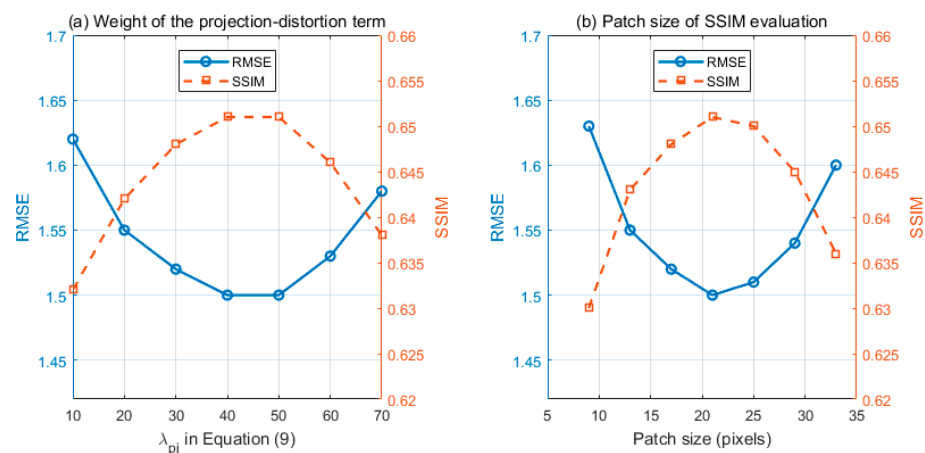


**Figure 12.** Sensitivity analysis of the proposed method with respect to (a) the weight  $\lambda_{pj}$  in Equation (9) and (b) the patch size of the SSIM-based quality evaluation. The remaining parameters are fixed at their default values. The left axis denotes RMSE and the right axis denotes SSIM.

## 5. Conclusions

Aiming at the three major challenges of parallax image stitching, including unstable feature points, geometric distortion, and seam artifacts summarized in the research motivation, this paper proposes a robust parallax image stitching framework based on content-preserving warping and quality-aware seam-cutting. By integrating dual-feature alignment, multi-term energy optimization, SSIM-based seam quality evaluation, SIFT flow local refinement, and adaptive seam integration, the proposed method effectively addresses the above technical challenges, eliminates artifacts, and maintains structural integrity. Sufficient quantitative and qualitative experiments covering multiple scenes and comparison methods verify the comprehensive performance and application potential of the proposed method in real parallax stitching scenarios. The framework is suitable for real-world applications such as panoramic imaging and UAV mapping.

**Author Contributions:** Conceptualization, J.W.; methodology, J.W. and W.W.; software, J.W.; validation, J.W. and S.Z.; data curation W.W. and S.Z.; supervision, G.C.; writing—original draft preparation, J.W.; writing—review and editing, J.W., W.W., G.C. and S.Z.; visualization, W.W.; funding acquisition J.W., W.W., and G.C.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported in part by the Project of Jiangxi Provincial Department of Education (Grant No. GJJ2402507).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The UDIS-D dataset [36] is publicly available. The remaining test image pairs used in this study are taken from the public datasets reported in references [12,14,20,25,31], as described in Section 4. The code used in this study is available from the corresponding author upon reasonable request.

**Acknowledgments:** The authors would like to express their sincere gratitude to Guobao Chen, for his invaluable guidance and constructive suggestions throughout the research and manuscript preparation. The authors also thank all members of the research group for their helpful discussions and technical support during data collection and experiments. Their valuable advice has greatly improved the quality of this study.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Li, Y.; Wang, J.; Yao, K. Modified phase correlation algorithm for image registration based on pyramid. *Alex. Eng. J.* **2022**, *61*, 709–718. [[CrossRef](#)]
2. Yan, X.; Cao, Y.; Yang, Y. Multi-modal image registration based on phase exponent differences of the Gaussian pyramid. *Remote Sens.* **2023**, *15*, 5764. [[CrossRef](#)]
3. Muhlich, J.L.; Chen, Y.A.; Yapp, C.; Russell, D.; Santagata, S.; Sorger, P.K. Stitching and registering highly multiplexed whole-slide images of tissues and tumors using ASHLAR. *Bioinformatics* **2022**, *38*, 4613–4621. [[CrossRef](#)] [[PubMed](#)]
4. Wang, J.; Ma, M.; Yan, S.; Zhang, J. Image stitching using double features-based global similarity constraint and improved seam-cutting. *Multimed. Tools Appl.* **2024**, *83*, 7363–7378. [[CrossRef](#)]
5. Dai, Q.; Fang, F.; Li, J.; Zhang, G.; Zhou, A. Edge-guided composition network for image stitching. *Pattern Recognit.* **2021**, *118*, 108019. [[CrossRef](#)]
6. Peng, Z.; Ma, Y.; Zhang, Y.; Li, H.; Fan, F.; Mei, X. Seamless UAV hyperspectral image stitching using optimal seamline detection via graph cuts. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–13. [[CrossRef](#)]
7. Ma, M.; Wang, J. Image stitching using sigmoid function to create perception mask. *J. Intell. Fuzzy Syst.* **2023**, *45*, 2047–2061. [[CrossRef](#)]
8. Gao, W.; Cui, C.; Liu, X.; Zhou, H. An improved image stitching method for blades of wind turbine based on online repair technology. *Comput. Ind.* **2026**, *176*, 104446. [[CrossRef](#)]

9. Huo, L.; Qin, L.; Wu, S.; Mi, S. Unsupervised confocal superficial eyelid image stitching: Flexible, accurate and smooth. *Biomed. Signal Process. Control* **2026**, *112*, 108760. [[CrossRef](#)]
10. Liu, F.; Gleicher, M.; Jin, H.; Agarwala, A. Content-preserving warps for 3D video stabilization. In *Seminal Graphics Papers: Pushing the Boundaries*; ACM: New York, NY, USA, 2023; Volume 2, pp. 631–639. [[CrossRef](#)]
11. Li, X.; He, L.; He, X. Combined regional homography-affine warp for image stitching. In *Proceedings of the 14th International Conference on Graphics and Image Processing (ICGIP 2022)*, Nanjing, China, 27–29 November 2022; SPIE: Bellingham, WA, USA, 2023; p. 127050U. [[CrossRef](#)]
12. Zaragoza, J.; Chin, T.J.; Brown, M.S.; Suter, D. As-projective-as-possible image stitching with moving DLT. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013*; IEEE: Piscataway, NJ, USA, 2013; pp. 2339–2346. [[CrossRef](#)]
13. Du, P.; Ning, J.; Cui, J.; Huang, S.; Wang, X.; Wang, J. Geometric structure preserving warp for natural image stitching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 21–24 June 2022*; IEEE: Piscataway, NJ, USA, 2022; pp. 3688–3696. [[CrossRef](#)]
14. Liao, T.; Li, N. Single-perspective warps in natural image stitching. *IEEE Trans. Image Process.* **2019**, *29*, 724–735. [[CrossRef](#)] [[PubMed](#)]
15. Xue, W.; Xie, W.; Zhang, Y.; Chen, S. Stable linear structures and seam measurements for parallax image stitching. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 253–261. [[CrossRef](#)]
16. Xiang, T.Z.; Xia, G.S.; Bai, X.; Zhang, L. Image stitching by line-guided local warping with global similarity constraint. *Pattern Recognit.* **2018**, *83*, 481–497. [[CrossRef](#)]
17. Lin, K.; Jiang, N.; Cheong, L.F. Seagull: Seam-guided local alignment for parallax-tolerant image stitching. In *Proceedings of the European Conference on Computer Vision (ECCV)*, Amsterdam, The Netherlands, 8–16 October 2016; Springer: Cham, Switzerland, 2016; pp. 370–385. [[CrossRef](#)]
18. Chen, J.; Li, Z.; Peng, C.; Mei, X. UAV image stitching based on optimal seam and half-projective warp. *Remote Sens.* **2022**, *14*, 1068. [[CrossRef](#)]
19. Su, Y.; Jiang, M.; Gao, L. Graph-cut-based collaborative node embeddings for hyperspectral images classification. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [[CrossRef](#)]
20. Li, N.; Liao, T.; Wang, C. Perception-based seam cutting for image stitching. *Signal Image Video Process.* **2018**, *12*, 967–974. [[CrossRef](#)]
21. Zhang, Z.; Wu, R.; Li, D. Image stitching and target perception for Autonomous Underwater Vehicle-collected side-scan sonar images. *Front. Mar. Sci.* **2024**, *11*, 1418113. [[CrossRef](#)]
22. Jia, Q.; Li, Z.J.; Fan, X. Leveraging line-point consistence to preserve structures for wide parallax image stitching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021*; IEEE: Piscataway, NJ, USA, 2021; pp. 12186–12195. [[CrossRef](#)]
23. Chen, X.; Yu, M.; Song, Y. Optimized seam-driven image stitching method based on scene depth information. *Electronics* **2022**, *11*, 1876. [[CrossRef](#)]
24. Liao, T.; Chen, J.; Xu, Y. Quality evaluation-based iterative seam estimation for image stitching. *Signal Image Video Process.* **2019**, *13*, 1199–1206. [[CrossRef](#)]
25. Liao, T.; Zhao, C.; Li, L. Leveraging local patch alignment to seam-cutting for large parallax image stitching. In *Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 11–18 October 2025*; IEEE: Piscataway, NJ, USA, 2025; pp. 27262–27271. [[CrossRef](#)]
26. Duan, H.; Min, X.; Sun, W. Attentive deep image quality assessment for omnidirectional stitching. *IEEE J. Sel. Top. Signal Process.* **2023**, *17*, 1150–1164. [[CrossRef](#)]
27. Cuturi, M. Soft-dtw: A differentiable loss function for time-series. In *Proceedings of the International Conference on Machine Learning (ICML)*, Sydney, Australia, 6–11 August 2017; PMLR: Cambridge, MA, USA, 2017; pp. 894–903. Available online: <https://proceedings.mlr.press/v70/cuturi17a.html> (accessed on 13 September 2026).
28. Wang, L.; Koniusz, P. Uncertainty-DTW for time series and sequences. In *Proceedings of the European Conference on Computer Vision (ECCV)*, Tel-Aviv, Israel, 23–27 October 2022; Springer: Cham, Switzerland, 2022; pp. 176–195. [[CrossRef](#)] [[PubMed](#)]
29. Qin, C.; Ran, X.; Zhang, D. Unsupervised image stitching based on generative adversarial networks and feature frequency awareness algorithm. *Appl. Soft Comput.* **2025**, *154*, 113466. [[CrossRef](#)]
30. Wang, J.; Sun, Y.; Liu, C. Semantics-guided hybrid framework for image stitching. *J. Electron. Imaging* **2026**, *35*, 023020. [[CrossRef](#)]
31. Nie, L.; Lin, C.; Liao, K.; Liu, S.; Zhao, Y. Parallax-tolerant unsupervised deep image stitching. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, France, 2–6 October 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 7365–7374. [[CrossRef](#)]
32. Bakurov, I.; Buzzelli, M.; Schettini, R.; Castelli, M. Structural similarity index (SSIM) revisited: A data-driven approach. *Expert Syst. Appl.* **2022**, *189*, 116087. [[CrossRef](#)]

33. Jiao, S.; Li, X.; Lu, X. An improved Otsu method for image segmentation. In *Proceedings of the 8th International Conference on Signal Processing, Guilin, China, 16–20 November 2006*; IEEE: Piscataway, NJ, USA, 2006. [[CrossRef](#)]
34. Liu, C.; Yuen, J.; Torralba, A. SIFT flow: Dense correspondence across scenes and its applications. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 978–994. [[CrossRef](#)] [[PubMed](#)]
35. 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](#)] [[PubMed](#)]
36. Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [[CrossRef](#)]
37. Von Gioi, R.G.; Jakubowicz, J.; Morel, J.M.; Randall, G. A Fast Line Segment Detector with a False Detection Control. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 722–732. [[CrossRef](#)] [[PubMed](#)]

**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.