

RESEARCH ARTICLE

Structure Preservation and Seam Optimization for Parallax-Tolerant Image Stitching

SHAOPING WEN^{1,2}, XIAOLEI WANG^{1,2}, WEICHAO ZHANG^{1,2}, GUANJUN WANG^{1,2,3}, MENGXING HUANG^{1,2}, (Member, IEEE), AND BENGUO YU⁴

¹State Key Laboratory of Marine Resource Utilization in South China Sea, Hainan University, Haikou 570228, China

²School of Information and Communication Engineering, Hainan University, Haikou 570228, China

³Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430074, China

⁴School of Biomedical Information and Engineering, Hainan Medical University, Haikou 571199, China

Corresponding authors: Guanjun Wang (wangguanjun@hainanu.edu.cn), Mengxing Huang (huangmx09@hainanu.edu.cn), and Benguo Yu (yubg@hainmc.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 62175054, Grant 61865005, and Grant 61762033; in part by the Natural Science Foundation of Hainan Province under Grant 2019CXTD400, Grant 620RC554, and Grant 617079; in part by the Major Science and Technology Project of Haikou City under Grant 2021-002; in part by the Open Project Program of Wuhan National Laboratory for Optoelectronics under Grant 2020WNLOKF001; in part by the National Key Technology Support Program under Grant 2015BAH55F04 and Grant 2015BAH55F01; in part by the Major Science and Technology Project of Hainan Province under Grant ZDKJ2016015; and in part by the Scientific Research Staring Foundation of Hainan University under Grant KYQD(ZR)1882.

ABSTRACT Parallax processing has long been a significant and challenging task in image stitching. In this paper, we study a new hybrid warping model based on multi-homography and structure preservation to achieve accurate alignment of regions at different depths while preserving local and global image structures. The homographies of different depth regions are estimated by dividing matching feature pairs into multiple layers. Then, layered warping is performed by determining the spatial relationships between the image mesh and these multi-homography, and then refining the local and global structural distortions through mesh optimization. Four constraints are considered during the local optimization process, including the local alignment error, global alignment error, and similarity error. In addition, we explore and introduce collinear structures into an objective function as a constraint for mesh optimization warping, which can preserve salient line structures while alleviating distortions in nonoverlapping areas. Furthermore, we develop an optimal seam search method based on seam error evaluation to improve the quality of the seams. Experimental results demonstrate that compared to existing methods, the proposed algorithm presents more accurate stitching results for images with large parallax and preserves salient image structures, and outperforming the existing methods both qualitatively and quantitatively.

INDEX TERMS Image stitching, multi-homography, optimal seam search, parallax, structure preservation.

I. INTRODUCTION

Image stitching is an important technology widely used in the field of computer graphics and vision. Its aim is to generate a high-quality, ghosting-free and wide-view image from multiple smaller source images at the lowest computational cost [1]–[3]. Commercial tools for security monitoring, virtual reality (VR) cameras, and outdoor live-broadcast cameras seem able to achieve good results in generating panoramic images in this way. However, one important fact

is overlooked: the input images for these tools change only with rotation or panning of the camera; under this assumption, an affine transformation and a perspective transformation can be used finish the stitching work well. When complex structures or parallax problems are present in captured images, the stitching results of the previous methods will suffer from ghosting artifacts and misalignment. Such parallax-tolerant image stitching has always been a challenging task.

At present, the solutions to the parallax tolerance problem have the following characteristics. First, adaptive models are based on spatially varying warping [4]–[7]. In early work, the global homography computed by matching scale


The associate editor coordinating the review of this manuscript and approving it for publication was Gulistan Raja .



FIGURE 1. Comparison of the results of a representative spatially varying warping method (APAP) [4] and our method. (a): The input image [13]. (b): The stitching results of APAP, which exhibit not only visual distortion but also obvious ghosting. (c): The stitching result of our approach, in which the distortion and ghosting are eliminated.

invariant feature transform (SIFT) features [8] has been used to align the input images, and good results can be achieved when the input image satisfies certain assumptions, e.g., when the camera is positioned with a single optical center or when the images contain only one principal plane. To obtain more natural results in scenes dominated by two planes, Gao *et al.* proposed using a dual-homology model for image stitching [7]. However, it is difficult for the dual-homology model to adapt to the diversity in real scenes and to refine local alignment. Zaragoza *et al.* proposed the method of as-projective-as-possible (APAP) warping to achieve the local alignment of target images [4]; this method is forward looking in the development of spatially varying warps based on grids. However, these methods still face the problems of ghosting in overlapping regions and perspective distortion in non-overlapping regions, especially in the presence of large parallax (see Fig. 1-(b)). Jing *et al.* proposed a method for parallax image stitching based on an elastic warping model, which adaptively removes incorrect local matches by means of a Bayesian feature refinement model and then calculates the optimal homography to improve the image alignment quality [19]. However, problems of local structural distortion and different degrees of ghosting when using a grid-based model and direct deformation strategy to process parallax images of multiple depth layers. To alleviate projective distortions and refine local alignment, researchers have introduced constraint terms into the mesh warping process to obtain smooth transitions. Moreover, it has been confirmed that adding content-preserving constraints during the warping process is beneficial for local adjustment [9], [10]. Pan *et al.* proposed a method combining global 2D homography with a local warping method based on mesh optimization, which produces accurate stitching results for local structures at the same depth in a scene [34]. However, these measures cannot guarantee accurate alignment for scenes with different depths captured in large-parallax images and lack the ability to preserve the global image structure; thus, they can easily cause distortion of nonoverlapping regions. Recently, Zhang *et al.* proposed an approach for content preserving image stitching with regular boundary constraints [11]; in this approach,

through suitable analysis of an irregular boundary, an optimal piecewise rectangular boundary can be constructed to reduce unwanted distortions. However, these methods lack ghosting processing for the overlapping areas of large-parallax images.

Second, seam cutting [10], [12] refers to the process of finding seams in overlapping regions while minimizing costs and then stitching the input images together at the seam. Such methods show good performance in object shape preservation. However, relying on traditional seam cutting will lead to mutilated and overlapped of objects. A third approach is known as optimal homography [13]–[15]. Xu *et al.* the optimal homography matrix is obtained by analyzing feature point distribution and minimizing the registration error [14]. Lee *et al.* proposed a warping residual-based image stitching method [15]. First, the images are segmented into superpixels, and then, the feature point weights are calculated based on residual deformation to find the optimal homography for each superpixel. Our approach differs from optimal homography methods in that we propose a strategy of optimizing the spatially varying warping process from coarse to fine, which is also adopted in the evaluation of seams; these strategies achieve more precise alignment results and the searched seams can introduce fewer artifacts.

In this paper, we study a new hybrid warping model based on multi-homography and structure preservation. According to our observations, it is unrealistic to assume that images that need to be stitched together will be captured while only translating or rotating the camera in daily life. Usually, large-parallax images will contain multiple different scene depths. When global alignment methods [1], [2] are applied to such images, ghosting artifacts can easily be generated in overlapping regions, and perspective distortion will appear in nonoverlapping regions. To address these problems, the homographies of different depth regions are estimated by dividing the matching points in the source images into different depth layers. Then, layered warping is performed based on the spatial relationships between the image mesh and the multi-homography to achieve the precise alignment of regions at different depths. To alleviate the distortion caused by parallax, a new structure-preserving warping method is

introduced to preserve the local and global image structures, especially the line structure, in order to achieve accurate image alignment. The basic mathematical model of this method comes from Liu [9]. Finally, we develop an optimal seam search method for seam error evaluation based on traditional seam cutting methods, which not only reduces ghosting and misalignment in overlapping regions but also introduces fewer artifacts (see Fig. 1-(c)).

In summary, the main contributions of the proposed method can be summarized as follows:

- 1) To the best of our knowledge, we are the first to use a hybrid warping model of layered warping and structure preservation to improve warping quality from coarse to fine.
- 2) We propose a new structure-preserving warping method to preserve the local and global image structures during warping.
- 3) We develop an optimal seam search method based on seam error evaluations to improve the seam quality and reduce artifacts.

The rest of this paper is organized as follows: Related works on the stitching of images with large parallaxes are introduced in Sec. 2. The details of the proposed approach are described in Sec. 3. Experimental results and comparisons with other methods are shown in Sec. 4. Finally, the paper is concluded in Sec. 5.

II. RELATED WORK

In previous image stitching techniques, the input images have been aligned by estimating the global homography. If the optical centers are not identical or the scene is not a single plane of input images, these methods cannot guarantee local alignment in the stitched image. To solve this problem, researchers have proposed spatially varying warping methods, content preservation methods and seam cutting methods. However, these methods still have deficiencies in regard to parallax tolerance and image structure preservation. A detailed investigation of image stitching and alignment principles has been presented in [16]. In this section, we will review relevant work from four perspectives: homography alignment, spatially varying warping, shape- and structure-preserving warping and seam cutting.

A. HOMOGRAPHY ALIGNMENT

When certain assumptions are satisfied, aligning overlapping images by using global homography [1], [2] can also yield satisfactory alignment results. However, if the necessary assumptions are not satisfied, the stitched images will be seriously misaligned. Gao *et al.* proposed a dual-homography alignment model containing two dominant planes; in this model, each image is divided into a far plane and a ground plane, and the homography values of the two planes are calculated separately [7]. This method achieves good results when there are only two depth layers (planes). Li *et al.* proposed a multiplane alignment method based on superpixel

segmentation to overcome the problem of image misalignment [17], but failed to address the challenges presented by large-parallax images.

B. SPATIALLY VARYING WARPING

For more flexible handling of local alignment, a series of spatially varying warping methods have been proposed. Lin *et al.* introduced a smooth affine splice field instead of a global affine transformation and processed parallel lines more flexibly by extending them to multiple planes [18]. However, the local alignment result could not be refined. Zaragoza *et al.* proposed the APAP warping method for image stitching, in which the moving direct linear transformation (DLT) method is used to estimate the optimal homography of the image grid and edge error adjustment is performed through local feature weighting [4]. To alleviate the perspective distortion caused by the APAP method, Lin *et al.* proposed a method that combines local homography and global similarity transformations, in which the global similarity is used to alleviate perspective distortion in nonoverlapping regions [6]. Li *et al.* proposed a mosaic image method based on robust elastic warping that uses the analytical warping function to eliminate parallax errors [19]. Recently, Li *et al.* proposed a local-adaptive image alignment method based on triangular facet approximation and introduced planar triangulation and spherical triangulation strategies for image alignment [20]. Chen *et al.* devised a strategy for using an objective function and a global similarity prior (GSP) to specify the desired characteristics of warps [21]. Herrmann *et al.* proposed the use of multiple registrations to reduce errors in the image stitching of scenes with significant depth variations or object motion [22]. However, their method has difficulty preserving linear structures without corresponding relationships in large-parallax image stitching. In contrast, our method can preserve both local and global linear structures.

C. SHAPE- AND STRUCTURE-PRESERVING WARPING

Shape- and structure-preserving warping mainly aims to preserve the similarity of image structures in order to generate stitching results with natural shapes and structures. Chang *et al.* proposed a spatial combination of projective transformation and similarity transformation to extrapolate the projective transformation of an overlapping region to non-overlapping regions in order to correct the shape of the stitched image (preventing excessive skew), and reduce distortion in non-overlapping regions [23]. However, this approach is similar to APAP [4] because structural distortion occurs in the overlapping area. Liao *et al.* propose single-perspective warps (SPW) to protect linear structures while suppress distortions [24]. Lin *et al.* proposed a mesh-based photometric alignment method that combines the superior performance of dense photometric alignment with the efficiency of mesh-based image warping to improve content-preserving robustness in low textured areas [25]. Luo *et al.* used positional relationship constraints on feature points and lines to accomplish accurate alignment [26]. These methods

show excellent performance in preserving the shapes and linear structures of images but cannot explicitly handle parallax.

D. SEAM CUTTING

Gao *et al.* proposed a seam-driven image stitching strategy, in which the optimal homography is selected based on a seam cutting loss and the seam with the lowest cost is selected as the location at which to stitch the input images [12]. Zhang *et al.* developed a stitching method that combines content-preserving warping (CPW) [9] and seam cutting; this method preserves the local shape of an image and reduces distortion [10]. Lin *et al.* guided the selection of the optimal homography by means of seam estimation to improve the stitching quality [13]. Li *et al.* proposed a visual perception-based seam cutting method in which the weight ratio of the color difference and saliency of different objects in the human visual system is simulated to optimize the energy function [27]. The traditional energy-minimizing seam evaluation method in large-parallax image stitching not only causes ghosting and misalignment in overlapping regions but also introduces artifacts.

III. THE PROPOSED APPROACH

In this section, we introduce our proposed method in reference to the common stitching process. The outline of our proposed method is shown in Fig. 2. Specifically, our hybrid warping model consists of two steps: first, matching points are divided into different depth layers to estimate the homographies of different depth regions, and then, layered warping is performed by determining the spatial relationships between the image mesh and multi-homography to achieve the precise alignment of regions at different depths. Subsequently, the mesh warping is optimized by means of structural constraints on the local alignment error, global alignment error, local similarity error, and line collinearity error. After completing the coarse-to-fine warping process, we apply an optimal seam search method based on seam error evaluation to find the optimal line at which to stitch the aligned images together. Our main contribution is that we propose a new image alignment and stitching method, which not only is the difficulty of aligning scenes with different depth layers in large-parallax images using global homography solved by means of a hybrid warping model, but also the salient line structure is also effectively preserved by introducing a line collinearity constraint. Moreover, an optimal seam search method yields image stitching results with minimal artifacts and ghosting.

A. MULTIHOMOGRAPHY ESTIMATION AND LAYERED WARPING

1) MULTIHOMOGRAPHY ESTIMATION

Given two partially overlapping images I and I' , matching feature points are obtained using SIFT [8], and the matching feature pairs are denoted by $\{P_\varphi, P'_\varphi\}_{\varphi=1}^n$. The projective transformation between the matching feature pairs can be

Algorithm 1 Local Homography Estimation

Input: Initial pair set $F_1 = F = \{P_\varphi, P'_\varphi\}_{\varphi=1}^n$, threshold N_{\min} and iteration index $k = 0$;

Output: The matching pair set T_k for each layer and the corresponding homography H_k for each layer;

```

1: repeat
2:    $k = k + 1$ ;
3:   Apply RANSAC to the matching pair set  $F_k$  for the
      model  $0_{3 \times 1} = P'_{k,\varphi} + HP_{k,\varphi}$ , where  $(P'_{k,\varphi}, P_{k,\varphi}) \in F_k$ ;
4:   Divide the inliers  $T'_{in}$  and outliers  $T'_{out}$  according to  $H$ ;
5:   if  $|T'_{in}| \geq N_{\min}$  then
6:     Set the matching pair set for the  $k$ -th layer to  $T_k$ ,
        $|T_k| = T'_{in}$ ;
7:     Set the homography of the  $k$ -th layer to  $H_k$ ,  $H_k = H$ ;
8:   end if
9:   Set a new matching pair set  $F_{k+1}$ , where  $F_{k+1} = T'_{out}$ ;
10: until  $|F_{k+1}| < N_{\min}$ .

```

expressed as $P' = H_g P$. The specific equation details are as follows:

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

where $H_g \in \mathbb{R}^{3 \times 3}$ is the global homography; (x', y') and (x, y) are the homogeneous coordinates of the matching feature points p' and p , respectively; and the value of h_{33} is 1.

Using the DLT [28] method, (1) can be rewritten as:

$$0_{3 \times 1} = ah = \begin{pmatrix} 0_{1 \times 3} & -P^T & y'P^T \\ P^T & 0_{1 \times 3} & -x'P^T \\ -y'P^T & x'P^T & 0_{1 \times 3} \end{pmatrix} \begin{pmatrix} h_{11} \\ \vdots \\ h_{33} \end{pmatrix}, \quad (2)$$

where the two rows of $a \in \mathbb{R}^{3 \times 9}$ are linearly independent. A constant h is set, and the quantity $\|a_i h\|$ is the algebraic error of the i -th datum. The sum of the squared algebraic errors is minimized, as follows:

$$\hat{h} = \arg \min_h \sum_{i=1}^n \|a_i h\|^2 = \arg \min_h \|Ah\|^2, \quad (3)$$

where $\|h\| = 1$ and the matrix $A \in \mathbb{R}^{2N \times 9}$ is obtained by vertically stacking the a_i for all $[1, n]$. The global homography H_g can be extrapolated from \hat{h} by calculating the least significant right singular vector of $A \in \mathbb{R}^{2N \times 9}$.

In this section, we group the features to compute the local homographies for different scene depth layers. First, the feature points are divided into different layers using the random sample consensus (RANSAC) algorithm [29], and then, the homography of each layer is estimated. The set of matching feature matching pairs in the k -th layer is denoted by F_k , the number of matching pairs in the k -th layer is $|F_k|$, and the corresponding homography of the k -th layer is H_k . To limit the warping similarity of each layer, we set a threshold N_{\min} , which represents the minimum number of matching pairs that $|F_k|$ must satisfy; otherwise, the process

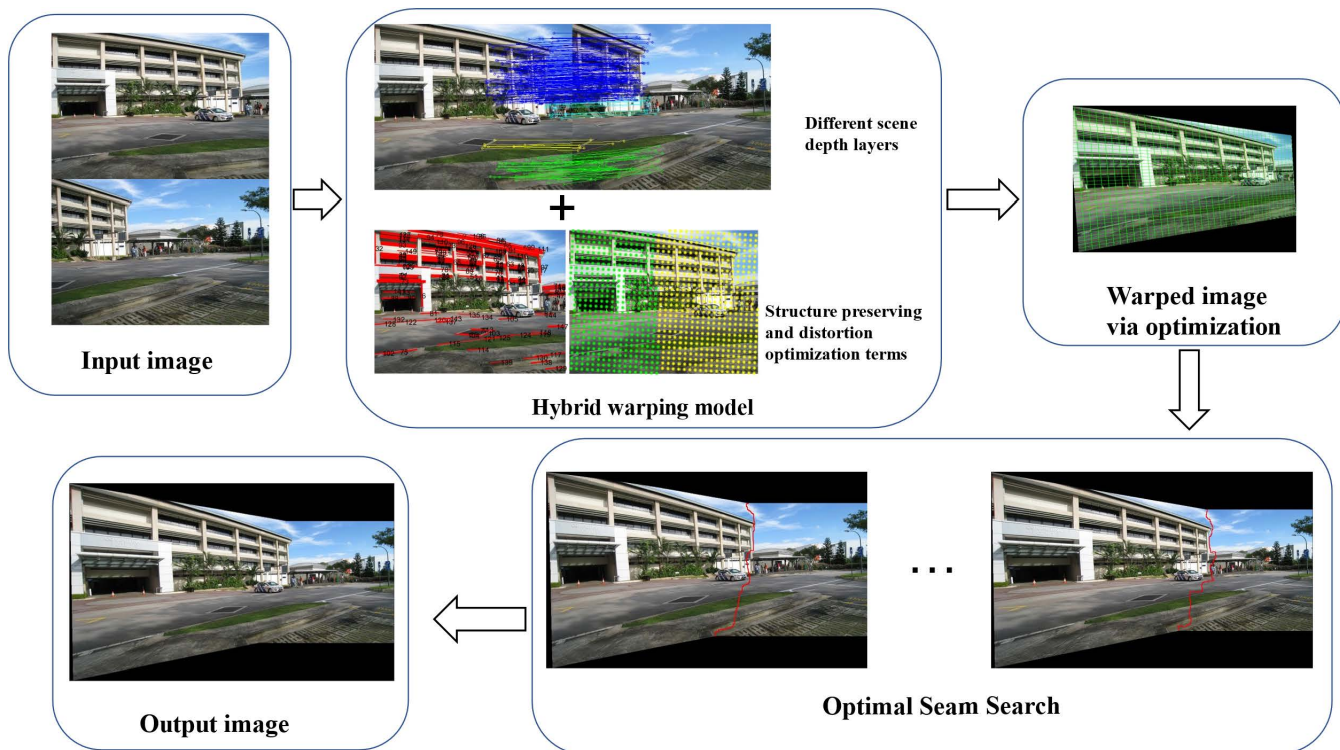


FIGURE 2. Summary of the proposed method. The hybrid warping model is used to accurately align and preserve the image structure, while the optimal seam search based on seam error evaluation is conducted to search for the optimal seam.

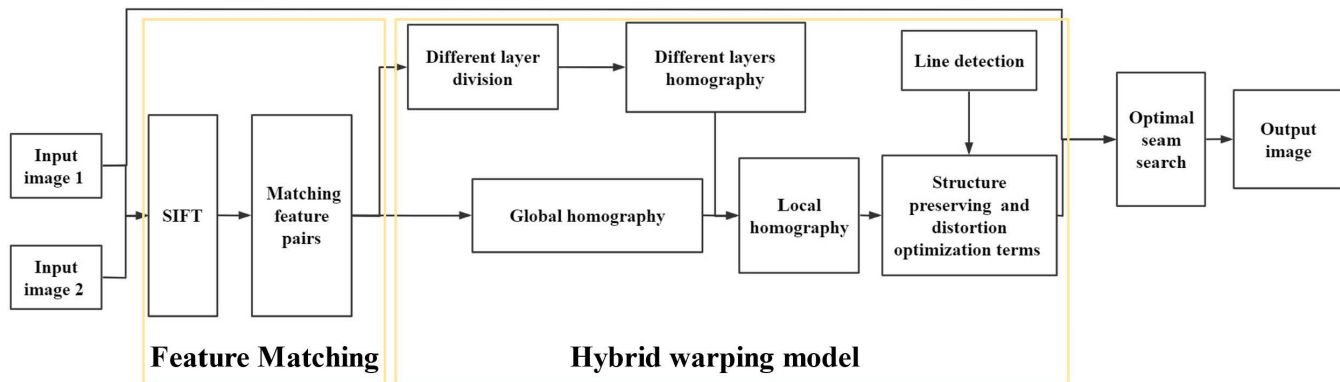


FIGURE 3. Flowchart of the proposed image stitching approach.

will be terminated at the k -th layer. The specific process of calculating the local homography corresponding to each layer is illustrated in Algorithm 1.

2) LAYERED WARPING

We combine the local homography H_k of each layer with the global homography H_g to obtain our layered warping result. First, we divide the target image into a unified mesh of $M \times N$ to improve our computing efficiency. Since arbitrary feature points usually do not coincide with any mesh vertices, the degree of dependence of the mesh warping on the local homographies of the different layers is determined in accordance with the distance between the mesh center and the nearest matching points of different feature layers.

The center of mesh cell g_i is represented by c_i , where i represents the specific i -th mesh cell, and the homography of mesh cell g_i is represented by H_i . Under the assumption that the feature points are finally separated into γ different layers, the final homography of mesh cell i is expressed as:

$$H_i = \sum_{k=2}^{\gamma} \omega_k^i H_k + (1 - \sum_{k=2}^{\gamma} \omega_k^i) H_g, \tag{4}$$

where ω_k^i is the weight of each layer of homography in the specified mesh. $\omega_k^i = d_k^i / \sum_{\varphi} d_{\varphi}^i$, where d_k^i is a position-dependent Gaussian weight

$$d_k^i = \max_{\varphi \in F_k} \exp\left(\frac{-\|c_i - p_{\varphi}\|^2}{\sigma^2}\right), \tag{5}$$

with σ being a scale constant.

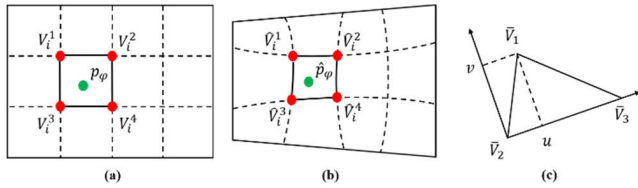


FIGURE 4. (a) the p_φ in the feature matching pair (p_φ, p'_φ) is represented by the bilinear interpolation weight of its four enclosing vertices. (b) \hat{p}_φ in the warped target image is represented in the same way. (c) Local coordinates of \hat{V}_1 in the triangle formed by \bar{V}_1, \bar{V}_2 and \bar{V}_3 .

B. STRUCTURE-PRESERVING WARPING

The multi-homography method addresses the problem of the alignment of image regions at different scene depths. However, structural dislocation may still exist in small local areas, which especially salient line structures. Liu *et al.* applied a content-preserving warping (CPW) method in the video stabilization work [9] and achieved satisfying results in local areas. Inspired by this success, we use CPW to further align target images pre-warped by multi-homography method. However, in contrast to CPW, our proposed structure-preserving warping method introduces a global alignment term and a line collinearity term to alleviate perspective distortion in large-parallax images and preserve line structure alignment. In our work, we use I, \bar{I} and \hat{I} respectively to represent the target images in three different stages—the input, the pre-warping result and the final warping result and still use the mesh divided on the target image in the previous stage to guide the local warping of the image. The mesh vertices in I, \bar{I} and \hat{I} are denoted by V_i, \bar{V}_i and \hat{V}_i , respectively. Any point p in the target image mesh can be represented by a linear combination of four mesh vertices: $V_i(P) = w_i V_i$, where $V_i = [V_i^1, V_i^2, V_i^3, V_i^4]$ denotes the four vertices of the mesh and $w_i = [w_i^1, w_i^2, w_i^3, w_i^4]^T$ denotes the weights assigned to each mesh vertices in the bilinear interpolation calculation. Thus, the point warping problem can be transformed into a mesh warping problem based on the optimization of an objective function. The energy terms contained in our proposed objective function are specified as follows:

1) LOCAL ALIGNMENT TERM E_p

Since feature points usually do not coincide with any mesh vertices, a point p_φ can be represented by a 2D bilinear interpolation of the four vertices of the smallest mesh cell enclosing it (see Fig. 4-(a)). For the warped vertices $\hat{V}_i = [\hat{V}_i^1, \hat{V}_i^2, \hat{V}_i^3, \hat{V}_i^4]$, we expect to use the same weights to represent the corresponding point \hat{p}_φ after warping (see Fig. 4-(b)). To align the feature points p_φ and \hat{p}_φ , we define the local alignment term to minimize the following distance:

$$E_p = \sum_{\varphi} \chi_{\varphi} \|w_i \hat{V}_i - \hat{p}_{\varphi}\|^2, \quad (6)$$

where n is the size of the matching feature set and χ_{φ} is the weight of the corresponding feature point.

2) GLOBAL ALIGNMENT TERM E_g

To force regions without feature points to be closer to the pre-warping result \bar{I} , we define the following global alignment terms to constrain excessive movement of the mesh vertices in nonoverlapping regions,

$$E_g = \sum_i \tau_i \|\hat{V}_i - \bar{V}_i\|^2, \quad (7)$$

where \bar{V}_i and \hat{V}_i are the corresponding mesh vertices in the pre-warping and structure-preserving warping results. τ_i is a binary value. We set this binary value to 1 if there is a feature point or an extracted line structure in the neighborhood of V_i ; otherwise, it is set to 0.

3) SMOOTHNESS TERM E_s

We use the similarity transformation constraint from CPW [9] to preserve the similarity of each mesh transformation while reducing the local distortion of the image during the warping process. The similarity transformation constraint measures the deviation of the similarity transformation between each twisted mesh element and its reference input mesh element. Each mesh cell is divided into two triangles. As shown in Fig. 4-(c), for each triangle $\Delta \bar{V}_1 \bar{V}_2 \bar{V}_3$, the specific position of vertex \bar{V}_1 in the local coordinate system can be represented in terms of vertices \bar{V}_2, \bar{V}_3 as follows:

$$\begin{aligned} \bar{V}_1 &= \bar{V}_2 + u(\bar{V}_3 - \bar{V}_2) + vR_{90}(\bar{V}_3 - \bar{V}_2), \\ R_{90} &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \end{aligned} \quad (8)$$

where (u, v) denotes the local coordinates of \bar{V}_1 . The local coordinates of the new vertex \hat{V}_1 are still represented in terms of \hat{V}_2, \hat{V}_3 . We want the deformation of each element in the mesh to approximate a similarity transformation. Therefore, we can guide the similarity transformation of a given triangle by minimizing the following similarity transformation cost term while preserving the spatial smoothness of the mesh cell warping.

$$\begin{aligned} E_s &= \sum_i \sum_{j=0}^8 w_s \|\hat{V}_{1,j} \\ &\quad - (\hat{V}_{2,j} + u(\hat{V}_{3,j} - \hat{V}_{2,j}) + vR_{90}(\hat{V}_{3,j} - \hat{V}_{2,j}))\|^2, \end{aligned} \quad (9)$$

where u and v are calculated in accordance with (8). and w_s is the saliency value of the triangle; we use the method in CPW [9] to measure the saliency of the triangle. Note that there are a total of four quadrangles containing \hat{V}_1 vertexes, and each quad has two such triangles. We consider all eight triangles containing vertex \hat{V}_1 and apply the corresponding similarity transformation to each triangle.

4) LINE COLLINEARITY TERM E_l

The process of mesh warping destroys the overall structure of a line. Therefore, it is necessary to add a line collinearity energy term to preserve the collinearity of the points on a line

after warping. Since long lines can be divided into short lines, the tension of long lines in an image can be enhanced by using the collinearity of corresponding short lines. First, we use the linear segment detector (LSD) [30] to detect line segments and then remove shorter lines. Let S_l be the set of straight-line segments in image I . We assign the labels $\{s^{l1}, s^{l2}, s^{l3} \dots s^{ln}\}$ to represent the endpoints of the short line segments into which the straight-line segment s^l is divided, and any s^{ly1}, s^{ly2} and s^{ly3} satisfy the following line segment proportions:

$$r^{ly} = \frac{|s^{ly2} - s^{ly1}|}{|s^{ly3} - s^{ly1}|}, \quad (10)$$

where $r^{ly} \in (0, 1)$. To preserve the proportion and orientation of the collinear line segments after the final warping, we define the line collinearity term as follows:

$$E_l = \sum_{l=1}^{|S_l|} \sum_y \left\| (\hat{s}^{ly2} - \hat{s}^{ly1}) - r^{ly} (\hat{s}^{ly3} - \hat{s}^{ly1}) \right\|^2, \quad (11)$$

where $|S_l|$ represents the number of straight-line segments in the set S_l and y is over all three-point combinations for the warped line in \hat{I} . The squared difference represents the distance between the different end points of the scaled warped collinear segment.

5) OBJECTIVE FUNCTION E

Based on the above four energy terms, the objective function is constructed as follows:

$$E = E_p + \lambda_1 E_g + \lambda_2 E_s + \lambda_3 E_l, \quad (12)$$

where λ_1, λ_2 and λ_3 are weight factors that control the contribution of each energy term ($\lambda_1 = 1, \lambda_2 = 0.5$ and $\lambda_3 = 10$ in our implementation). The problem of minimizing the above objective function can be solved by a sparse linear solver.

C. OPTIMAL SEAM SEARCH

The seam cutting method is widely used to alleviate the artifacts caused by local dislocations of parallax images. Existing seam quality assessment methods are usually formulated in terms of energy minimization. However, in large-parallax image stitching, seam quality assessment methods that completely rely on energy minimization cannot find an optimal seam consistent with human visual perception. Therefore, this paper introduces an optimal seam search method based on a seam error evaluation to improve seam quality and reduce the parallax in overlapping areas. The details of the method are as follows.

We use N to represent the overlapping area between images I and I' , define a label set $L = \{0, 1\}$, and assign a label $l_p \in L$ to each pixel p_i of the seam within the overlapping area, where 0 corresponds to I and 1 corresponds to I' . The goal of traditional seam cutting methods is to find a set of labels l (i.e., a mapping from p to L) that minimizes the following energy function:

$$E(l) = \sum_{p \in N} D_p(l_p) + \sum_{(p,q) \in Z} S_{p,q}(l_p, l_q), \quad (13)$$

where $Z \subset N \times N$ is the collection of neighborhood pixels, the data term $D_p(l_p)$ represents the penalty term for assigning label l_p to a pixel $p \in N$, and the seam smoothing term $S_{p,q}(l_p, l_q)$ represents the cost of assigning a pair of labels (l_p, l_q) to a pair of pixels (p, q) . The seam smoothing term is defined as:

$$S_{p,q}(l_p, l_q) = \frac{1}{2} |l_p - l_q| (I_*(p) + I_*(q)), \quad (14)$$

$$I_*(p) = \|I(\cdot) - I'(\cdot)\|_2 \quad (15)$$

where $I_*(\cdot)$ denotes the Euclidean-metric difference map. Please refer to [31] for more details about the minimization solution for the energy function in (13).

The goal of the hybrid warping model is to achieve the precise alignment of different depth regions while preserving local and global structural similarity. Then, we compute the structural similarity index (SSIM) score [32] and zero mean normalized cross-correlation (ZNCC) score to evaluate the pixel errors along the seams. We define a 17×17 local patch centered on the pixel of the seam and define the local patch cost of the pixel as:

$$E_{patch}(p_i) = \frac{2 - (\text{SSIM}(p_i) + \kappa \text{ZNCC}(p_i))}{4}, \quad (16)$$

where the index value range of SSIM and ZNCC is $[-1, 1]$, and their values are 1 only when the two patches are exactly the same. κ is a weighting factor ($\kappa = 0.35$).

For large-parallax images capturing different scene depths, relying on the local patch cost to evaluate seam pixel errors at the edges of different depth layers may result in erroneous evaluations. Thus, point cost terms are added along the seam pixels to refine the seam error estimation.

$$E_{point}(p_i) = \frac{\|I(p_i) - I'(p_i)\| + \|I(q_i) - I'(q_i)\|}{2}, \quad (17)$$

where p_i and q_i represent adjacent pixels in the overlapping region. The point cost term is used to measure the difference between the pixels on both sides of the seam to effectively solve the problem of incorrect evaluation of local patches. Experimentally, it is found that this analysis based on local patches endows the local image with better ‘‘continuity,’’ and the point cost term plays an important role in the pixel evaluation at edges in the image. The patch cost and point cost have the same monotonicity when significant errors are produced by the developed seam error assessment method:

$$E(p_i) = \zeta \cdot \hat{E}_{patch}(p_i) \cdot \hat{E}_{point}(p_i), \quad (18)$$

where \hat{E} is the denoised costs and ζ is an error scale preservation coefficient. With the above seam error evaluation method, we adopt an iterative updating approach to find the optimal seam. Large seam errors usually indicate misaligned pixels. Thus, in the optimal seam search process, we adjust the smoothing cost of misaligned pixels at the seam by modifying the function $f(x) = e^{\sigma(x-\varepsilon)}$. Specifically, we define the following function to calculate the update smoothing term:

$$\tilde{I}_d(p_i) = \begin{cases} f(E(p_i)) \cdot I_d(p), & p \in N(\text{seam}), \\ I_d(p), & \text{otherwise.} \end{cases} \quad (19)$$



FIGURE 5. The datasets used for the ablation experiments in this paper, from left to right are the “Fence,” “Sidewalk,” and “House” images from Lin [6].

where N (seam) is the strip area obtained by extending the left and right sides of the current seam by 5 pixels. Then, the modified function is used to estimate the seam energy. This procedure is iterated multiple times until currently estimated seam is completely contained within the previous seam strip region; then, the search for a new seam is terminated, and the current seam is considered the optimal seam.

Fig. 3 illustrates the flow of the proposed method and the related algorithms. The hybrid warping model consists of two parts: including coarse warping, which considers different scene depth layers and fine warping, in which the mesh changes are optimized using structure-preserving constraints. In the coarse warping stage, the matching feature points are first divided into different depth layers; then, the homographies of the different layers are obtained; and finally uses formula (4) to calculate the local homography to obtain the pre-warping image \tilde{I} . The objective function for structure-preserving optimization is composed of local alignment error, global alignment error, local similarity error, and line collinearity error terms. We regard the image warping problem as a mesh warping problem, which is defined as an optimization problem. The goal is to align the pre-warping image \tilde{I} to the reference image while avoiding significant structural distortion. The representation illustrated in Fig. 4 is used to adjust the relationships between each mesh vertex and corresponding mesh points through the structure preserving optimization item, and thus, the final warping image \hat{I} is obtained. Finally, the image is stitched by iteratively searching for the optimal seam to minimize the seam error to reduce artifacts in the overlapping area, which is difficult to achieve with pixel fusion methods.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

We used 18 pairs of test images to evaluate the performance of the proposed image stitching algorithm. The dataset sources include image data published in [6], [10], [22], [33], and the data can be found on corresponding project websites or in their supplementary materials. In addition, three real-world image sets were captured by our group. We compared the performance of the proposed algorithm with that of five existing algorithms: APAP [4], ELA [19], GSP [21], SPHP [23] and SPW [24]. The parameters of the algorithms considered for comparison were set in accordance with the suggestions in the original papers. The experiments were conducted on a PC with an Intel I7-11800h 2.3 GHz CPU and 16 GB of RAM.

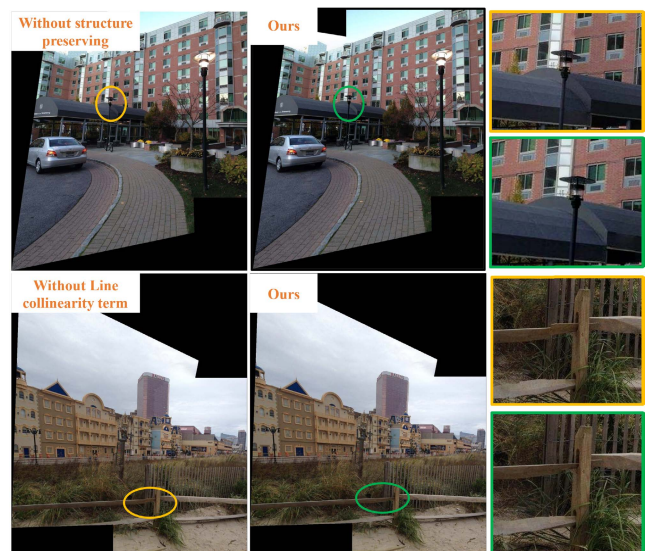


FIGURE 6. Ablation study on the “Sidewalk,” and “Fence” images. The first column from top to bottom shows the results of our method without the structure-preserving term (top), and our method without the line collinearity term (bottom). The second column shows the results of our full method. The yellow boxes show the distortion and misalignment caused by the lack of structure-preserving warping, and the green boxes show images with a more natural and reasonable appearance.

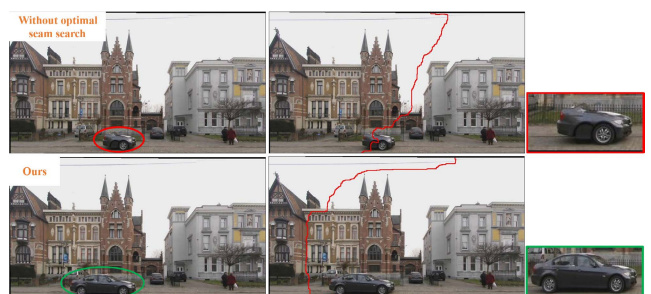


FIGURE 7. Results of the ablation experiment on the “Street” image. The first row shows the results of our method without the optimal seam search, where the red box shows the clipping and ghosting caused by parallax. The second row shows the result of our optimal seam stitching method, where the green box shows the same image region as the red box but with a more natural and reasonable appearance.

A. ABLATION STUDY

To better test and evaluate the performance of each component of the algorithm, we conducted an ablation experiment. In this experiment test images, including the “Fence,” “Sidewalk,” and “House” images from Lin [6] (see Fig. 5). The results verify the unique contributions of the structure-preserving warping, the linear collinearity term and the optimal seam search. More details are shown in Fig. 6 and Fig. 7. We summarize the main results of the ablation experiments as follows:

- 1) Structure-preserving warping refines local alignment and preserves salient scene structures during warping.
- 2) The line collinearity term is key for preserving the overall structure of lines.
- 3) The optimal seam search method based on the seam error evaluation improves the quality of the seams and introduces fewer artifacts.

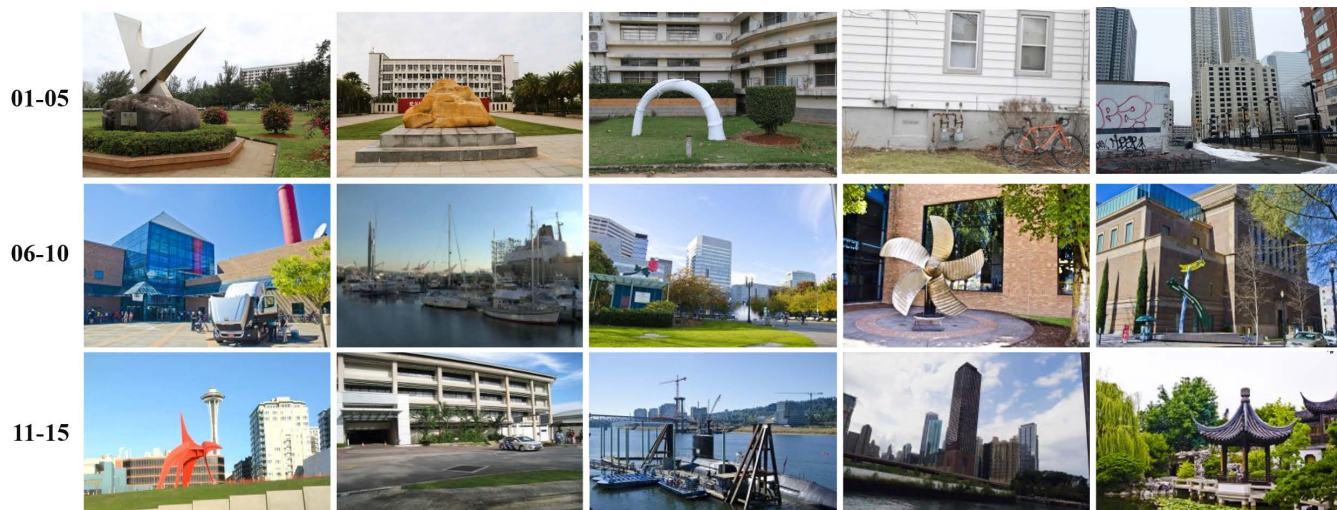


FIGURE 8. The dataset used in the comparison experiments in this paper. The 01-03 image data pairs (“Landmark,” “Granite,” and “Water pipe”) were captured by our group, the 04-05 image data pairs (“Bike” and “Graffiti building”) from [33] and [22], and the comparative experiment on the 01-05 data groups demonstrates the advantages and disadvantages of each algorithm in detail. The 06-15 image data pairs from [10], from top to bottom, are the “Truck,” “Boat,” “Lawn,” “Propeller,” “Building,” “Seattle,” “Carpark,” “Submarine,” “Building group,” and “Garden” images, and these data are used for a secondary proof.

As shown in Fig. 6, the layered warping process can align image regions at different depth layers but cannot constrain the local shape warping results in the images; e.g., without structure-preserving warping, the tops of the streetlights in the “Sidewalk” image suffer from obvious distortion. Second, the introduction of the line collinearity term into the structure-preserving warping process helps preserve the overall structure of lines in large-parallax images. For example, without line collinearity constraint, the foreground objects in the “Fence” image suffer from significant misalignment. We conclude that the proposed hybrid warping model can better align local objects at different scene depths and preserve structural stability when dealing with large-parallax images (the second column). As shown in Fig. 7, when there is partial occlusion in the foreground of the large-parallax image, our optimal seam search still finds the visually perceptual optimal seam (second row) and alleviates the artifacts introduced by traditional seam cutting methods and other seam driven methods (first row). These comparative experiments show that our method preserves salient scene structures during warping and produces high-quality seams.

B. COMPARISON WITH EXISTING METHODS

To further evaluate the performance of our proposed method, we compare it with five existing algorithms, namely APAP [4], ELA [19], GSP [21], SPHP [23] and SPW [24]. APAP uses spatial distance weighting and the moving DLT method to calculate the warping of the image mesh cells. ELA adaptively aligns grid cells by using analytical warping functions and Bayesian feature refinement models. GSP utilizes a grid-based local alignment model and a global similarity prior to control image warping. SPHP extrapolates the projective transformation from an overlapping region to at nonoverlapping region by combining the projection transform and the similarity transform in space to reduce the distortion in

nonoverlapping regions. SPW uses single-perspective warps to preserve linear structures while suppress distortion.

Among the experimental images (see Fig. 8), the image data pairs 01–03 (“Landmark,” “Granite,” and “Water pipe”) were collected by our group; image data pairs 04 and 05 (“Bike” and “Graffiti building”) were obtained from [33] and [22], respectively; and image data pairs 06–15 were obtained from [10]. We used the first 5 pairs of images as the first group of test data and the rest as the second group of test data. The scene depth characteristics of the first group (01–05) of test images are not complicated, but the parallax caused by the variability of the capture positions and angles makes image warping and fusion challenging. As seen in Fig. 9, these comparative experiments show that existing methods cannot accurately align foreground objects at different scene depths and produce severe ghosting artifacts in overlapping areas, such as the silver object in the “Landmark” image, the granite in the “Granite” image, the water pipe object in the “Water pipe” image and the graffiti pattern in “Graffiti building” image. In addition, although the window in the background of the “Bike” image is well aligned by the existing methods, the bicycle wheels in the stitched image are incorrectly aligned. Moreover, the background regions in both the “Granite” image stitched with GSP and the “Graffiti building” images stitched with ELA and GSP are distorted.

In contrast, to reduce the distortion of the background objects in each image, our method takes into account different scene depth layers through layered warping and adaptively optimizes the local warping through structure-preserving warping. Finally, we apply our optimal seam search method to improve the seam quality and reduce the ghosting artifacts caused by parallax effects. As shown in Fig. 9, our method results in natural and reasonable appearance for the granite object in the “Granite” image, the water pipe object in the “Water pipe” image and the graffiti pattern in the “Graffiti

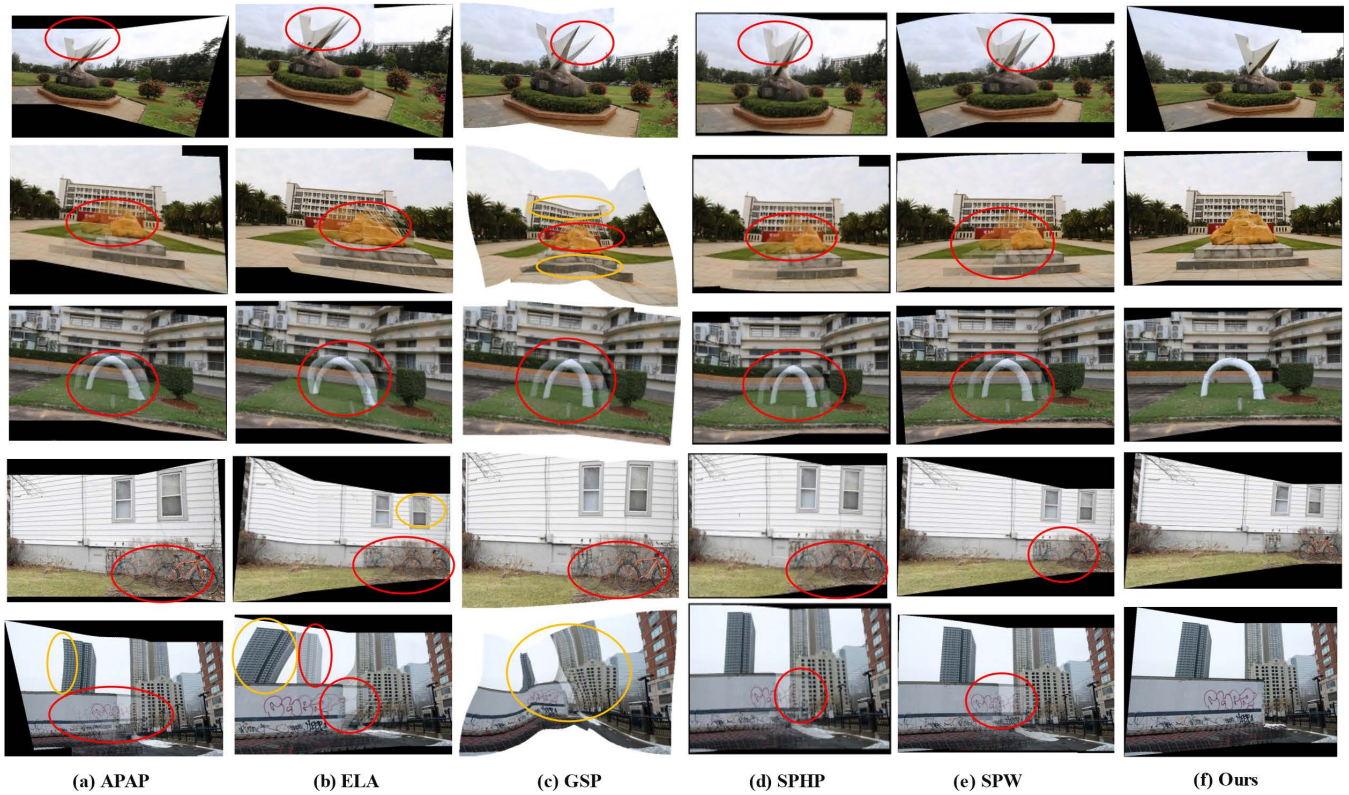


FIGURE 9. The results of our method on the first group of test data for comparison with those of the five existing methods: APAP [4], ELA [19], GSP [21], SPHP [23] and SPW [24]. From top to bottom are the “Landmark,” “Granite,” “Water pipe,” “Bike,” and “Graffiti building” images are shown.

TABLE 1. Comparison on PSNR and SSIM.

dataset	APAP		ELA		GSP		SPHP		SPW		OURS	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Truck	28.8192	0.5436	17.9038	0.4625	19.1708	0.8565	21.3667	0.7044	20.7437	0.8665	29.8386	0.9054
Boat	17.4230	0.4618	26.1370	0.7194	26.0707	0.7309	12.9855	0.2057	22.0920	0.5558	26.5365	0.8818
Lawn	28.2211	0.7594	25.9022	0.8077	28.0172	0.8302	27.8620	0.8046	24.4188	0.6564	28.6155	0.8688
Propeller	27.4236	0.7006	26.4711	0.6125	27.8524	0.6546	28.0648	0.8207	28.2624	0.8595	28.3873	0.8619
Building	23.3947	0.5248	28.4614	0.6443	28.6997	0.8009	28.5356	0.8111	28.2125	0.7915	28.9070	0.8492
Seattle	21.3451	0.5954	13.7759	0.4546	19.0647	0.6542	20.7973	0.6401	22.9637	0.7242	27.7183	0.9627
Carpark	26.3930	0.9096	20.9056	0.8183	24.9202	0.7073	25.1509	0.7618	26.1535	0.9203	26.4422	0.9493
Submarine	27.8756	0.8956	25.6649	0.6008	28.0108	0.9094	26.5340	0.7965	26.2397	0.7245	28.4893	0.9269
Building group	19.2723	0.6333	12.0164	0.5657	13.3840	0.5275	13.5816	0.5846	14.6650	0.5421	27.7839	0.8221
Garden	25.8612	0.7605	22.8354	0.6543	23.9849	0.8037	23.0754	0.7944	26.9377	0.8606	28.5160	0.8993

building” image. The warping of background objects, such as the background buildings in the “Graffiti building” and “Granite” images, also looks more natural.

The images in the second group of test data (06–15) contain various objects at different depths of field, which makes it more challenging to obtain natural and aesthetically pleasing stitching results. The results of the corresponding comparative experiments are shown in Fig. 10. Due to the

small parallax in the “Propeller” and “Carpark” images, the existing methods can produce acceptable results. However, because the pixels of some foreground objects in the reference image and the target image do not satisfy the requirements of the neighborhood operation, these foreground objects at different scene depths exhibit different degrees of ghosting artifacts when the existing methods are used. Examples include the truck object in the “Truck” image, the flower



FIGURE 10. The results of our method on the second group of test data for comparison with those of the five existing methods: APAP [4], ELA [19], GSP [21], SPHP [23] and SPW [24]. From top to bottom are the “Truck,” “Boat,” “Lawn,” “Propeller,” “Building,” “Seattle,” “Carpark,” “Submarine,” “Building group,” and “Garden” images are shown.

region of the “Lawn” image, the umbrella object in the “Building” image, the red object in the “Seattle” image, the car in the “Carpark” image, the iron bracket region in the “Submarine” image, the top of the skyscraper in the “Building group” image, and the top of the pavilion in the “Garden” image.

In addition, the line structures are distorted in the mesh warping method, e.g., the mast in the “Boat” image and the pillar part in the “Garden” image. On the other hand, our proposed hybrid warping model can align both local and global structures while the exploiting an optimal seam search to effectively avoid ghosting artifacts on foreground objects, e.g., the red object in the “Seattle” image and the top of the skyscraper in the “Building group” image.

Furthermore, we compute the SSIM score [32] and PSNR score to quantitatively evaluate the stitching quality of the

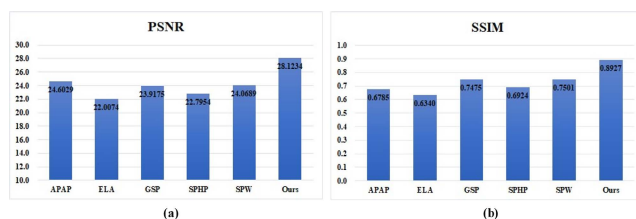


FIGURE 11. Quantitative performance comparison of the image stitching results of the different algorithms. (a): Average PSNR. (b): Average SSIM.

different methods: the results are shown in Table 1. Considering the diversity of the test data, we also compute the mean values to compare the SSIM and PSNR scores. The average PSNR scores obtained by APAP, ELA, GSP, SPHP, SPW, and our method are 24.6029, 22.0074, 23.9175, 22.7954, 24.0689 and 28.1234, respectively (see Fig. 11-(a)), and the

average SSIM scores are 0.6785, 0.6340, 0.7475, 0.6924, 0.7501 and 0.8927 (see Fig. 11-(b)). These results show that our algorithm achieves satisfactory scores in terms of both the SSIM and PSNR metrics.

V. CONCLUSION

In this paper, we study a new hybrid warping model based on multi-homography and structure preservation. Layered warping is performed based on the spatial relationship between the image mesh and the multi-homography to achieve precise alignment of image regions at different depths. Both local and global structures are preserved during warping by structure-preserving warping. Finally, we apply an optimal seam search method based on seam error evaluation to improve the seam quality. On test images with challenging degrees of parallax, we compared the results of our method with those of the APAP, ELA, GSP, SPHP and SPW methods. The experimental results show that the method achieves accurate stitching results for large parallax images and outperform some state-of-the-art warping algorithms in both qualitative and quantitative aspects. However, there is still no clear improvement in the stitching efficiency. In future work, we will incorporate deep learning to optimize the alignment of different depth layers in imaged scenes. We will also conduct a study of curvilinear structural preservation using linear features. Moreover, we will investigate the use of an end-to-end image stitching network to improve processing efficiency.

REFERENCES

- [1] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *Int. J. Comput. Vis.*, vol. 74, no. 1, pp. 59–73, Aug. 2007.
- [2] J. D. Mehta and S. G. Bhirud, "Image stitching techniques," in *Thinkquest*. New Delhi, India: Springer, 2010, pp. 74–80.
- [3] W. Yan, C. Hou, J. Lei, Y. Fang, Z. Gu, and N. Ling, "Stereoscopic image stitching based on a hybrid warping model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 9, pp. 1934–1946, Sep. 2017.
- [4] J. Zaragoza, T.-J. Chin, M. S. Brown, and D. Suter, "As-projective-as-possible image stitching with moving DLT," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2339–2346.
- [5] J. Li, P. Jiang, S. Song, H. Xia, and M. Jiang, "As-aligned-as-possible image stitching based on deviation-corrected warping with global similarity constraints," *IEEE Access*, vol. 7, pp. 156603–156611, 2019.
- [6] C.-C. Lin, S. U. Pankanti, K. N. Ramamurthy, and A. Y. Aravkin, "Adaptive as-natural-as-possible image stitching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1155–1163.
- [7] J. Gao, S. J. Kim, and M. S. Brown, "Constructing image panoramas using dual-homography warping," in *Proc. CVPR*, Jun. 2011, pp. 49–56.
- [8] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, pp. 91–110, Dec. 2004.
- [9] F. Liu, M. Gleicher, H. Jin, and A. Agarwala, "Content-preserving warps for 3D video stabilization," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 1–9, 2009.
- [10] F. Zhang and F. Liu, "Parallax-tolerant image stitching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 3262–3269.
- [11] Y. Zhang, Y.-K. Lai, and F.-L. Zhang, "Content-preserving image stitching with piecewise rectangular boundary constraints," *IEEE Trans. Vis. Comput. Graphics*, vol. 27, no. 7, pp. 3198–3212, Jul. 2021.
- [12] J. Gao, Y. Li, T. J. Chin, and M. S. Brown, "Seam-driven image stitching," in *Proc. Eurographics (Short Papers)*, 2013, pp. 45–48.
- [13] K. Lin, N. Jiang, L. F. Cheong, M. Do, and J. Lu, "SEAGULL: Seam-guided local alignment for parallax-tolerant image stitching," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 370–385.
- [14] P. Xu, K. Chen, A. Xie, and S. Xiong, "Parallax-tolerant image stitching with optimal homography," in *Proc. J. Phys., Conf.*, 2019, vol. 1237, no. 3, Art. no. 032034.
- [15] K.-Y. Lee and J.-Y. Sim, "Warping residual based image stitching for large parallax," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 8198–8206.
- [16] R. Szeliski, "Image alignment and stitching: A tutorial," *Found. Trends Comput. Graph. Vis.*, vol. 2, no. 1, pp. 1–104, 2006.
- [17] J. Li, D. Wu, P. Jiang, Z. Li, and S. Song, "Locally aligned image stitching based on multi-feature and super-pixel segmentation with plane protection," *IEEE Access*, vol. 9, pp. 168315–168328, 2021.
- [18] W.-Y. Lin, S. Liu, Y. Matsushita, T.-T. Ng, and L.-F. Cheong, "Smoothly varying affine stitching," in *Proc. CVPR*, Jun. 2011, pp. 345–352.
- [19] J. Li, Z. Wang, S. Lai, Y. Zhai, and M. Zhang, "Parallax-tolerant image stitching based on robust elastic warping," *IEEE Trans. Multimedia*, vol. 20, no. 7, pp. 1672–1687, Jul. 2018.
- [20] J. Li, B. Deng, R. Tang, Z. Wang, and Y. Yan, "Local-adaptive image alignment based on triangular facet approximation," *IEEE Trans. Image Process.*, vol. 29, pp. 2356–2369, 2020.
- [21] Y.-S. Chen and Y.-Y. Chuang, "Natural image stitching with the global similarity prior," in *Proc. Eur. Conf. Comput. Vis.* Springer, Cham, 2016, pp. 186–201.
- [22] C. Herrmann, C. Wang, R. S. Bowen, E. Keyder, M. Krainin, C. Liu, and R. Zabih, "Robust image stitching with multiple registrations," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 53–67.
- [23] C.-H. Chang, Y. Sato, and Y.-Y. Chuang, "Shape-preserving half-projective warps for image stitching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 3254–3261.
- [24] T. Liao and N. Li, "Single-perspective warps in natural image stitching," *IEEE Trans. Image Process.*, vol. 29, pp. 724–735, 2020.
- [25] K. Lin, N. Jiang, S. Liu, L.-F. Cheong, M. Do, and J. Lu, "Direct photometric alignment by mesh deformation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2405–2413.
- [26] X. Luo, Y. Li, J. Yan, and X. Guan, "Image stitching with positional relationship constraints of feature points and lines," *Pattern Recognit. Lett.*, vol. 135, pp. 431–440, Jul. 2020.
- [27] N. Li, T. Liao, and C. Wang, "Perception-based seam cutting for image stitching," *Signal, Image Video Process.*, vol. 12, no. 5, pp. 967–974, Jul. 2018.
- [28] Z. Zhang, "Parameter estimation techniques: A tutorial with application to conic fitting," *Image Vis. Comput.*, vol. 15, no. 1, pp. 59–76, 1997.
- [29] M. A. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [30] R. G. von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, "LSD: A fast line segment detector with a false detection control," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 4, pp. 722–732, Apr. 2010.
- [31] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.
- [32] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [33] C. Herrmann, C. Wang, R. S. Bowen, E. Keyder, and R. Zabih, "Object-centered image stitching," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 821–835.
- [34] X. Pan and G. Wang, "Parallax-tolerant image stitching based on mesh optimization," in *Proc. IEEE 2nd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Mar. 2017, pp. 414–420.



SHAOPING WEN is currently pursuing the M.S. degree with the College of Information and Communication Engineering, Hainan University, Haikou, China. His research interests include image stitching, video stitching, computer graphic, and computer vision.



XIAOLEI WANG is currently pursuing the M.S. degree with the College of Information and Communication Engineering, Hainan University, Haikou, China. Her research interests include fiber-optic sensors, Fabry–Perot interferometers, and digital image processing.



WEICHAO ZHANG received the B.S. degree from Yangzhou University, Yangzhou, China, in 2019. He is currently pursuing the M.S. degree with Hainan University, Haikou, China. His current research interests include computer vision, video understanding, and image processing.



GUANJUN WANG received the B.S. degree in electronic science and technology and the M.S. degree in physical electronics from Zhengzhou University, Zhengzhou, China, in 2005 and 2008, respectively, and the Ph.D. degree in optical engineering from Beihang University, Beijing, China, in 2012. From 2013 to 2015, he was with the Institute of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province, Shenzhen University, Shenzhen, China, as a Postdoctoral Researcher. Then, he was with the North University of China, Taiyuan, China, as an Associate Professor. In 2018, he joined Hainan University. He is currently an Associate Professor. He has published more than 30 academic articles as the first or corresponding author. His research interests include optical fiber sensor and optical interrogation.



MENGXING HUANG (Member, IEEE) received the Ph.D. degree from Northwestern Polytechnical University, in 2007. Then, he joined the Staff with the Research Institute of Information Technology, Tsinghua University, as a Postdoctoral Researcher. In 2009, he joined Hainan University. He is currently a Professor and a Ph.D. Supervisor in information and communication engineering, and the Dean of the School of Information Science and Technology. He is also the Executive Vice-President of the State Key Laboratory of Marine Resource Utilization in the South China Sea, and the Director of the Hainan Key Laboratory of Big Data and Smart Services. He has published more than 60 academic articles as the first or corresponding author. He has reported 12 patents of invention, owns three software copyright, and published two monographs and two translations. His current research interests include big data and intelligent information processing, ocean information perception and fusion AI, and smart service. He is a Senior Member of Chinese Computer Federation (CCF), and a member of the Information System Committee in CCF. He has awarded two Second Class and one Third Class Prizes of The Hainan Provincial Scientific and Technological Progress.



BENGUO YU received the M.S. degree in technology of computer application and the Ph.D. degree in measurement technology and instruments from the North University of China, Taiyuan, China, in 2006 and 2013, respectively. He went to study at York University, Canada, from December 2011 to January 2013. In 2019, he joined Hainan Medical University. He is currently an Associate Professor. He has published more than 14 academic articles as the first or corresponding author, and published three monographs. His research interests include big data and artificial intelligence, data analysis, and educational informatization.

• • •