

Stereoscopic Image Stitching Based on a Hybrid Warping Model

Weiying Yan, Chunping Hou, Jianjun Lei, *Member, IEEE*, Yuming Fang, *Member, IEEE*,
Zhouye Gu, *Member, IEEE*, and Nam Ling, *Fellow, IEEE*

Abstract—Traditional image editing techniques cannot be directly used to process stereoscopic media, as extra constraints are required to ensure consistent changes between the left and right images. In this paper, we propose a hybrid warping model for stereoscopic image stitching by combining projective and content-preserving warping. First, a uniform homography algorithm is proposed to prewarp the left and right images, and thus ensure consistent changes. Second, a content-preserving warping is introduced to locally refine alignment and reduce vertical disparities. Finally, a seam-cutting-based algorithm is used to find a blending seam, and the multiband blending algorithm is used to produce the final stitched image. Experimental results show that the proposed method can effectively stitch stereoscopic images, which not only avoids local distortions, but also reduces vertical disparities reasonably.

Index Terms—Image stitching, image warping, stereoscopic image.

I. INTRODUCTION

IMAGE stitching has been widely studied in the fields of computer vision and graphics. It can be used to stitch images of adjacent views with small horizon fields into a single image with large horizon. At present, image stitching techniques are integrated in many well-known commercial softwares, such as Adobe Photoshop, AutoStitch, and so on. These softwares are widely used by the users to organize and process photo collections.

Traditionally, image stitching is typically implemented by using a parametric projective warping to bring images into alignment [1]–[3]. The projective warping method is often implemented by a homography matrix. Although 2D projective warping is robust, it can only provide accurate alignment

for planar scenes or parallax-free camera motions, i.e., the photographer's physical location must be fixed and only rotational motion is permitted. In other cases, 2D projective warping results in artifacts, such as ghosting or broken structures. Recent image stitching methods adopt spatially varying warping [4]–[6] to align input images, which produces fewer artifacts than 2D projective warping. However, these methods are primarily designed for 2D images.

With the rising popularity of stereoscopic images and videos, techniques for producing and editing stereoscopic media have recently attracted a large amount of attention in the research community. For stereoscopic image stitching, a limited number of studies have been reported in [7]–[11]. For example, Zhu *et al.* [10] presented a method to automatically and efficiently generate stereoscopic mosaics using a video camera mounted on an airborne platform. Gurrieri and Dubois [11] proposed two multiple-camera configurations capable to produce high-quality stereoscopic panoramas in real time. However, these methods might need to densely sample the scene using a video camera and/or follow some specific rules to rotate the camera. In addition, they do not work well with a sparse set of captured input images.

Compared with traditional 2D media content, stereoscopic media provides a more immersive viewing experience by producing a perception of depth. With this additional depth dimension, more challenges and constraints emerge in creating enjoyable 3D experiences [12]–[14]. Since naive extensions of the existing 2D stitching algorithms usually fail to deliver a comfortable 3D viewing experience, new techniques are required to accommodate these new constraints and to take advantage of new information provided by stereoscopic media. Fig. 1 shows some of the issues that arise when producing stereoscopic content using a 2D image stitching method, as-projective-as-possible algorithm (APAP) [4]. Although no ghost exists in both the left and right images processed by the APAP method, the sizes of the left and right images are different. Moreover, the disparity of the resulting images is not consistent, which will result in discomfort and eye fatigue. This problem is primarily caused by different projective models used in the left and right images, and traditional warp estimation without considering the vertical disparity.

In this paper, we present a novel stitching method for stereoscopic images. The contributions of this paper are summarized as follows: 1) we provide an effective stereoscopic image stitching method based on a hybrid warping model, which combines projective and content-preserving warping; 2) we propose a uniform homography matrix to simultaneously prealign neighboring stereoscopic images, which glob-

Manuscript received July 25, 2015; revised October 16, 2015 and March 13, 2016; accepted April 27, 2016. Date of publication May 6, 2016; date of current version September 5, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61471262, in part by the Natural Science Foundation Key International (Regional) Cooperation Research Projects under Grant 61520106002, in part by the Ph.D. Programs Foundation within the Ministry of Education of China under Grant 20130032110010, and in part by the China Scholarship Council. This paper was recommended by Associate Editor C. Zhang. (*Corresponding author: Jianjun Lei.*)

W. Yan, C. Hou, and J. Lei are with the School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China (e-mail: jilei@tju.edu.cn).

Y. Fang is with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: fa0001ng@e.ntu.edu.sg).

Z. Gu is with ARRIS Group Inc., San Diego, CA 92121, USA (e-mail: guzh0001@gmail.com).

N. Ling is with the Department of Computer Engineering, Santa Clara University, Santa Clara, CA 95053 USA (e-mail: nling@scu.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSVT.2016.2564838

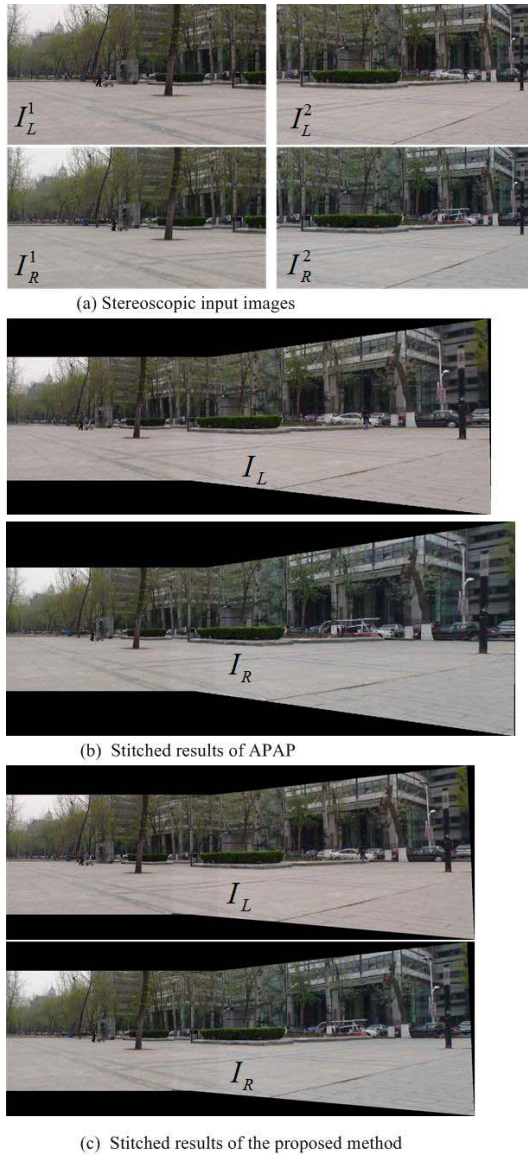


Fig. 1. Result samples from different stereoscopic image stitching algorithms.

ally preserves the structure consistency of adjacent images; 3) a content-preserving method is introduced and solved by an energy minimization solution to locally refine prewarped stereoscopic images and reduce vertical disparities; and 4) an extended seam-cutting-based algorithm is designed to find a seam of blending, and the multiband blending algorithm is accordingly used to produce the final stitched images.

The rest of this paper is organized as follows. Section II gives a brief overview on the existing 2D image stitching methods as well as the recent work on warping-based image editing techniques. Section III presents the proposed stereoscopic image stitching method. The experimental results are shown in Section IV. Finally, this paper is concluded in Section V.

II. RELATED WORK

A. Image Stitching

In general, there are two steps in traditional image stitching methods, namely, registration and blending. Registration

techniques try to align two input images to the same coordinate system according to the extracted geometrical motion model, while blending techniques, such as seam cutting [15], [16] and multiblending [17], [18], combine the two registered images into a single one. The seam-cutting methods [15], [16] optimize pixel selection among the overlapping images to minimize visible seams, and the advanced pixel blending techniques, such as Laplacian pyramid blending [17] and Poisson image blending [18], stitch two registered images into a single composite image.

In order to reduce ghosting artifacts, registration has to be performed before blending for image stitching. To register two images, most existing image stitching methods estimate a 2D projective transformation or homography between two input images and use it to align them [1]–[3]. A homography transformation H between two input images I^1 and I^2 can be described as

$$\begin{bmatrix} w \cdot x_1 \\ w \cdot y_1 \\ w \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} \quad (1)$$

where $(x_1, y_1, 1)$ and $(x_2, y_2, 1)$ are the homogenous coordinates of two points in image I^1 and I^2 , respectively. However, for actual complex scenes, these homography-based methods cannot adequately characterize the required warp, and thus lead to ghosting. Recent spatially varying warping methods, including APAP warping methods [4], smoothly varying affine (SVA) stitching [5], and shape-preserving half-projective warps for image stitching [6], have been extended to image stitching and work well. Zaragoza *et al.* [4] proposed a bundle adjustment technique to simultaneously refine multiple APAP warps for large panoramas. Lin *et al.* [5] developed an SVA stitching field, which is flexible enough to handle parallax while allowing local deformations. Chang *et al.* [6] designed a spatial combination method using a projective transformation and a similarity transformation to align the input image and preserve the perspective of individual images as similarity warps. These methods reduce ghosts under the condition of keeping the geometric realism of perspective image stitching. However, all these methods are designed for 2D images. When these methods are used for stereoscopic image stitching, they lead to an inconsistent global structure and incorrect disparities when stitching the left and right images, as introduced in Section I.

B. Image Editing Based on Warping

Warping techniques have been widely used in many aspects of stereoscopic image editing. Early image warping techniques are designed to address various problems in 2D image/video processing, including the correction of wide-angle lens distortion [20], [21], the production of projections for artistic perspective [22], and so on. In [20], a content-based approach was presented to map a wide-angle photograph defined on the viewing sphere onto a flat image, such that the resulting image appears more natural. Wei *et al.* [21] proposed a framework for fisheye video correction, which minimizes time-varying distortion and preserves salient content in a coherent manner. Carroll *et al.* [22] addressed the problem of manipulating

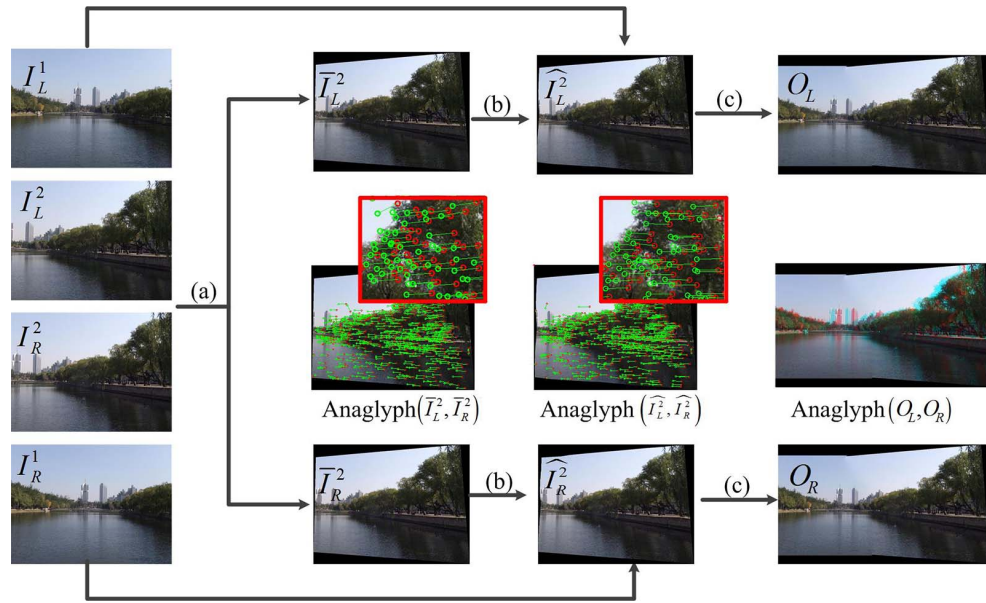


Fig. 2. Outline of the proposed method. (a) Prewarping based on a uniform homography. (b) Image fine-tuning based on content-preserving warping. (c) Image blending.

perspective by using a 2D warping method guided by constraints based on projective geometry. While, inspired by the widely accepted saliency theory [23], [27], content-aware warping methods have been applied to many aspects of image processing, including image resizing [28], [29], video compression [30], [31], and image classification [22]. The key idea of content-aware warping techniques is that, as long as the warped results of salient regions are closer to the physically correct constraints, the distortions introduced in nonsalient regions might be unnoticeable to the observer. Recently, warping methods are also expanded to 3D image/video processing with stereoscopic constraints, including content-aware 3D image resizing [33], [34], 3D video stabilization [35], stereoscopic image cloning and pasting [36]–[38], the manipulation of disparity for stereoscopic images [39], [40], and changing perspective for stereoscopic images [41]. Chang *et al.* [33] proposed a content-preserving warp method for stereoscopic image retargeting, which mainly preserves salient regions during image resizing by building some constraint conditions to compute a warped image. Liu *et al.* [35] proposed a content-preserving warp based on sparse correspondences in monocular image sequences. Tong *et al.* [38] proposed a stereoscopic image composition method, which can efficiently paste 2D images into stereoscopic scenes by warping them appropriately, which generate pleasant results for viewers. Niu *et al.* [42] extended a traditional 2D image warping technique for stereoscopic images. Their method can construct a good stereoscopic image whose structure is different from the original scene structure. Lang *et al.* [39] proposed an effective nonlinear disparity mapping method, which deforms input stereoscopic video streams to meet a desired disparity range based on image warping. Du *et al.* [41] addressed the problem of manipulating perspective in stereoscopic pairs and applied their method to stereoscopic stitching. However, their method cannot ensure the alignment of the overlapping region of two stereoscopic images, since it only considers

the changes in image perspective. Zhang and Liu [43] first stitched the left images of input stereoscopic images and input disparity maps, and then warped and stitched the right images constrained by the stitched left image and target disparity map. Their method could align the left and right images without ghosts and keep the consistency of the stitched left and right images based on the target disparity map. However, the method may produce some distortions or vertical disparities, since they only warp the right images rather than both, to adjust disparities and keep the consistency.

In this paper, we propose a novel stitching method for a pair of stereoscopic images based on a hybrid warping model. First, we design a uniform homography matrix to simultaneously prewarp neighboring stereoscopic images, which can globally preserve the structure consistency of the two images. Furthermore, inspired by the widely applied content-aware warping methods, we propose to use a content-aware method to fine-tune the prewarped stereoscopic images and minimize vertical disparities. Experimental results show that the proposed method can effectively stitch stereoscopic images.

III. PROPOSED STEREOSCOPIC IMAGE STITCHING METHOD

Fig. 2 shows an outline of the proposed method, which mainly includes prewarping based on a uniform homography, image fine-tuning based on content-preserving warping, and image blending. Let I_L^1 , I_R^1 , I_L^2 , and I_R^2 represent the input stereoscopic image pairs, and I_L^1 and I_R^1 are chosen as the reference image pairs. First, we propose a homography matrix to roughly prealign the pair of input stereoscopic images I_L^2 and I_R^2 . Such a homography matrix can preserve the global structure consistency of two input images. The transformed images are shown as $\bar{I}_L^2$ and $\bar{I}_R^2$, and the corresponding anaglyph is represented as anaglyph($\bar{I}_L^2$ and $\bar{I}_R^2$). The green and red points in Fig. 2 are the

scale-invariant feature transform (SIFT) feature points in the warped left and right images, respectively. It can be observed from the enlarged portion of the anaglyph that there are still some vertical disparities after the processing of homography-based prealignment. Second, the content-preserving warping method is used to fine-tune the alignment and avoid visual distortion. The results after content-preserving warping are denoted by $\hat{I}_L^2$ and $\hat{I}_R^2$, and the corresponding anaglyph is represented as anaglyph ($\hat{I}_L^2$ and $\hat{I}_R^2$). Finally, a seam-cutting-based seam finding algorithm and the multiband blending algorithm [16] are deployed to create final stitching results, denoted by O_L and O_R .

A. Prewarping Based on the Uniform Homography Matrix

Applying 2D stitching methods to stereoscopic images may not work well due to the difficulty in ensuring consistent size and disparity. Therefore, we propose a uniform projective transformation to two stereoscopic images. The homography matrix H is estimated from the sets of matched SIFT feature points [44], which are across (I_L^1, I_L^2) , (I_R^1, I_R^2) , and (I_L^2, I_R^2) . All the matched feature points are verified using the fundamental matrix estimated by Random Sample Consensus [45]. The sets F^1 , F^2 , and F^3 of the matched feature points are denoted as

$$F^1 = \{(f_{il}^1, f_{il}^2) | f_{il}^1 \in I_L^1, f_{il}^2 \in I_L^2, i = 1 \dots n_1\} \quad (2)$$

$$F^2 = \{(f_{jr}^1, f_{jr}^2) | f_{jr}^1 \in I_R^1, f_{jr}^2 \in I_R^2, j = 1 \dots n_2\} \quad (3)$$

$$F^3 = \left\{ (f_{kl}^2, f_{kr}^2) \left| \begin{array}{l} f_{kl}^2 \in (I_L^2 - I_{OL}), \text{ and} \\ f_{kr}^2 \in (I_R^2 - I_{OR}), \quad k = 1 \dots n_3 \end{array} \right. \right\} \quad (4)$$

where (f_{il}^1, f_{il}^2) , (f_{jr}^1, f_{jr}^2) , and (f_{kl}^2, f_{kr}^2) denote the pairs of the matched feature points from (I_L^1, I_L^2) , (I_R^1, I_R^2) , and (I_L^2, I_R^2) , respectively. n_1 , n_2 , and n_3 denote the number of feature points. I_{OL} denotes the overlapping region of I_L^1 and I_L^2 , and I_{OR} denotes the overlapping region of I_R^1 and I_R^2 . To calculate the homography matrix H , the energy term E_{pre} is defined by considering alignment and the consistency of disparity, which can be expressed as

$$E_{pre} = \frac{1}{n_1} \sum_{i=1}^{n_1} \left\| \frac{1}{w_i} H f_{il}^2 - f_{il}^1 \right\|^2 + \frac{1}{n_2} \sum_{j=1}^{n_2} \left\| \frac{1}{w_j} H f_{jr}^2 - f_{jr}^1 \right\|^2 + \frac{1}{n_3} \sum_{k=1}^{n_3} \left| \left[\frac{1}{w_k} H f_{kl}^2 \right]_y - \left[\frac{1}{w_k} H f_{kr}^2 \right]_y \right|^2 \quad (5)$$

where H denotes a homography matrix with eight parameters. w_i , w_j , and w_k are the homogenous components of the corresponding feature points' coordinates. $[(1/w_k) H f_{kl}^2]_y$ and $[(1/w_k) H f_{kr}^2]_y$ are the y -coordinates of feature points after transformation.

In (5), H is obtained through three stages as follows.

- 1) Four pairs of feature points are randomly selected from F^1 , and a homography matrix H_1 is computed by (1). Then, H_1 is plugged into (5) to obtain the energy term E_{pre1} . Similarly, four pairs of feature points are also selected from F^2 , and a homography matrix H_2 is computed by (1). H_2 is plugged into (5) to obtain the energy term E_{pre2} .

- 2) E_{pre} is calculated as $E_{pre} = \min(E_{pre1}, E_{pre2}, E_{pre})$. Then, the corresponding H of E_{pre} can be obtained. Here, the initial value of E_{pre} is set as an infinity number. The corresponding initial H is set as a 3×3 unit matrix $E_{3 \times 3}$.
- 3) Go to step 1), iteratively until the iterative time n satisfies the predefined value. In our simulations, the predefined value is experimentally set as 1000.

With the above procedure, H bringing the minimum energy value E_{pre} in the iteration can be obtained. After obtaining H , the two images I_L^2 and I_R^2 are transformed to $\bar{I}_L^2$ and $\bar{I}_R^2$. Then, the two stereoscopic images are roughly aligned with the transformation.

B. Alignment Fine-Tuning and Disparity Consistency Adjustment Using Content-Preserving Warping

As indicated in Section III-A, after the prewarping based on uniform homography matrix, the two stereoscopic images are roughly aligned. However, there still exist some vertical disparities. Currently, various content-preserving warping methods have been used in many applications, such as video stabilization and image/video retargeting. Existing studies [33], [34] show that the content-warping method can be used for small local adjustment. Therefore, we use it to further align prewarping results, as shown in Fig. 2(b).

In the proposed warping method, the prewarping images $\bar{I}_L^2$ and $\bar{I}_R^2$ are divided into $m \times n$ grid cells. The vertices in $\bar{I}_L^2$ and $\bar{I}_R^2$ are denoted by $\bar{V}_{i,j,l}^2$ and $\bar{V}_{i,j,r}^2$. $\bar{F}^1$, $\bar{F}^2$, and $\bar{F}^3$ are the sets of matched feature points obtained in the image pairs $(\bar{I}_L^1, \bar{I}_L^2)$, $(\bar{I}_R^1, \bar{I}_R^2)$, and $(\bar{I}_L^2, \bar{I}_R^2)$, respectively.

1) *Sparse Stereoscopic Correspondences*: The resulting feature sets $\bar{F}^1$, $\bar{F}^2$, and $\bar{F}^3$ generally have an irregularly clustered distribution of correspondences, depending on scene content. Moreover, these feature sets may include mismatched feature points. To ensure a better spatial distribution of features and avoid mismatched points, a pruning algorithm is applied to remove correspondences around the vertices $\bar{V}_{i,j,l}^2$ and $\bar{V}_{i,j,r}^2$, and a convolution method is applied to remove the isolated correspondences points in featureless regions. The pruning algorithm removes all the pairs from $\bar{F}^1$, $\bar{F}^2$, and $\bar{F}^3$, which meet the conditions of

$$\| \bar{f}_{i,l}^2 - \bar{V}_{i,j,l}^2 \| < \delta \quad (6)$$

$$\| \bar{f}_{i,r}^2 - \bar{V}_{i,j,r}^2 \| < \delta \quad (7)$$

where $\bar{f}_{i,l}^2$ and $\bar{f}_{i,r}^2$ are feature points in the images $\bar{I}_L^2$ and $\bar{I}_R^2$, respectively. The radius δ depends on the image resolution and the size of the grid cell. In our experiments, δ is set as 10 empirically.

Apart from removing the irregular correspondences, the convolution method is used to remove the isolated correspondences in stereoscopic images. In certain grid cell with fewer points, there is no feature point in its eight neighborhood grid cells or just one neighborhood grid cell includes feature points. We call these feature points in the grid cell as isolated correspondences. To obtain a promising warping result, these isolated correspondences should be removed.

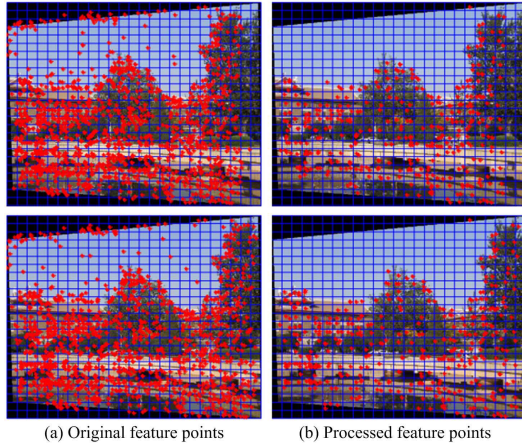


Fig. 3. Sparse stereoscopic correspondence.

The isolated correspondences are detected by the convolution method, which can be expressed as

$$G = M_{m \times n} * R_{3 \times 3} \quad (8)$$

where $M_{m \times n}$ is a matrix consisting of 0 and 1 (0 denotes that there is no feature point in a grid cell, while 1 denotes there are feature points in a grid cell). $R_{3 \times 3}$ denotes a 3×3 matrix with all the values as 1. The convolution method is separately used for the image $\hat{I}_L^2$ and $\hat{I}_R^2$. If the value of an element is 1 or 2 in the matrix G , the feature points and its correspondences in the corresponding grid cell are removed. Fig. 3 shows an example of the resulting features.

2) *Warping*: After the selection of sparse feature points, the image warping problem is formulated as a grid cell warping problem, which can be expressed as

$$\widehat{V}_{i,j}^2 = W(\overline{V}_{i,j}^2) \quad (9)$$

where $\overline{V}_{i,j}^2$ includes $\overline{V}_{i,j,l}^2$ and $\overline{V}_{i,j,r}^2$; the values of $\widehat{V}_{i,j}^2$ are unknown variables, which include $\widehat{V}_{i,j,l}^2$ and $\widehat{V}_{i,j,r}^2$. A feature point $\hat{f}$ in image $\hat{I}^2$ is denoted by a linear combination of four neighboring cell vertices of the corresponding cell. If the feature point $\hat{f}$ located in the mesh grid (i, j) , it can be formulated as

$$\hat{f} = \sum_{p=i}^{i+1} \sum_{q=j}^{j+1} a_{p,q} \widehat{V}_{p,q}^2 \quad (10)$$

where the coefficients $a_{p,q}$ are computed by finding $\overline{V}_{p,q}^2$ and inverting the bilinear interpolation process [46] from $\overline{V}_{p,q}^2$ to $\hat{f}$ in the image $\hat{I}^2$. The mesh warping problem is addressed by solving an energy minimization framework based on defined energy terms, including alignment, stereoscopic constraint, and shape preservation. The energy terms are described in detail as follows.

a) *Alignment term*: The alignment term is used to better align the prewarping images and the reference images, including local and global alignment terms. For local alignment, the feature points $\hat{f}_{i,l}^2$ and $\hat{f}_{i,r}^2$ should be moved to match their corresponding positions in the reference images I_L^1 and I_R^1 ,

so that they can be well aligned. Thus, the local alignment term is defined as

$$E_h = \sum_{f_{i,l}^1 \in \overline{F}^1} \|\hat{f}_{i,l}^2 - f_{i,l}^1\|^2 + \sum_{f_{i,r}^1 \in \overline{F}^2} \|\hat{f}_{i,r}^2 - f_{i,r}^1\|^2. \quad (11)$$

Since the feature points in sets $\overline{F}^1$ and $\overline{F}^2$ are from the overlapping image region, the local alignment term only constrains the warping of the overlapping image region with selected feature points. For other regions, as the prewarped results $\hat{I}_L^2$ and $\hat{I}_R^2$ have already provided a good approximation, our method encourages the regions without feature points to be as close to the prewarping result as follows:

$$E_g = \sum_{j=1}^n \sum_{i=1}^m g_{i,j,l} \|\widehat{V}_{i,j,l}^2 - \overline{V}_{i,j,l}^2\|^2 + \sum_{j=1}^n \sum_{i=1}^m g_{i,j,r} \|\widehat{V}_{i,j,r}^2 - \overline{V}_{i,j,r}^2\|^2 \quad (12)$$

where $g_{i,j,l}$ is a binary value. If there is no feature point in the grid cell of the left image, $g_{i,j,l} = 1$. Otherwise, it is equal to 0. $g_{i,j,r}$ is defined in the similar way for the right image.

b) *Stereoscopic constraint term*: The stereoscopic constraint term is designed to obtain comfortable 3D experiences, including vertical and horizontal disparity constraints. Our method enforces the stereoscopic constraint between the two images by applying the constraint to a sparse set of matching feature points, since the human vision system is sensitive to these features.

The vertical disparity constraint is used to ensure the vertical alignment of feature points, namely, the y-coordinates of the corresponding feature points in both the images should be similar. The energy term of vertical disparity constraint is defined as

$$E_y = \sum_{i=1}^{\overline{n}_3} \|\hat{f}_{i,l,y}^2 - \hat{f}_{i,r,y}^2\|^2 \quad (13)$$

where $\overline{n}_3$ is the number of feature points in the set $\overline{F}^3$.

The horizontal disparity constraint keeps the disparity consistency of the prewarping results $\hat{I}_L^2$ and $\hat{I}_R^2$, i.e., the x-coordinate of each feature point is encouraged to be close to the disparity of feature point in the prewarping image. The energy term of horizontal disparity constraint is described as

$$E_d = \sum_{i=1}^{\overline{n}_3} \|\hat{f}_{i,l,x}^2 - \hat{f}_{i,r,x}^2 - \varphi(d_i)\|^2 \quad (14)$$

where $\varphi(d_i) = \overline{f}_{i,l,x}^2 - \overline{f}_{i,r,x}^2$. In addition, our method can achieve the user-defined disparity distribution by changing the disparity mapping $\varphi(d_i)$. An example is shown in Section IV-B.

c) *Shape preservation term*: Shape preservation term is used to reduce distortion and keep the smoothness of warping. We would like to get the warping results, which are smooth and locally similar, since such warping results appear less distorted. In this case, the warping function $W(\overline{V}_{i,j}^2)$ should satisfy the Cauchy–Riemann equations, and the Hessian of the

warping function $W(\widehat{V_{i,j}^2})$ should be zero. Carroll *et al.* [22] converted these equations into energy function based on finite differences. The Cauchy–Riemann equation is equivalent to the energy function of similarity transformation [35], and the Hessian is equivalent to smoothness energy function. The energy function of similarity transformation and the smoothness energy function are, respectively, shown in (15) and (16). For the calculation of squared norm in (16), the matrix is first converted to a vector row-by-row as the work in [47] and [48] as

$$E_{s1}(i, j) = \sum_{(i,j) \in m_{ij}} \left\| \widehat{V_{i+1,j,l}^2} - \widehat{V_{i,j,l}^2} - u \cdot (\widehat{V_{i,j+1,l}^2} - \widehat{V_{i,j,l}^2}) - v \cdot R \cdot (\widehat{V_{i,j+1,l}^2} - \widehat{V_{i,j,l}^2}) \right\|^2 \quad (15)$$

$$E_{s2}(i, j) = \sum_{\substack{(i,j) \\ \in m_{ij}}} \left\| \begin{bmatrix} \widehat{V_{i,j+1,l}^2} - 2\widehat{V_{i,j,l}^2} + \widehat{V_{i,j-1,l}^2} \\ \widehat{V_{i+1,j,l}^2} - 2\widehat{V_{i,j,l}^2} + \widehat{V_{i-1,j,l}^2} \\ \widehat{V_{i+1,j+1,l}^2} - \widehat{V_{i+1,j,l}^2} - \widehat{V_{i,j+1,l}^2} + \widehat{V_{i,j,l}^2} \end{bmatrix} \right\|^2 \quad (16)$$

where $R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$; m_{ij} denotes a grid cell; u and v are the known coordinates within the local coordinate system. Since our grid cell is rectangle, we set u as 0 and $v = \|\widehat{V_{i+1,j}^2} - \widehat{V_{i,j}^2}\| / \|\widehat{V_{i,j+1}^2} - \widehat{V_{i,j}^2}\|$.

In our method, the salient regions are encouraged to undergo more restrictive smoothing than the nonsalient ones. Thus, the shape preservation term of the left image is defined as

$$E_{s1} = \sum_{j=1}^n \sum_{i=1}^m \omega_{(i,j)} \cdot (E_{s1}(i, j) + E_{s2}(i, j)) \quad (17)$$

where $\omega_{(i,j)}$ is the saliency value of the grid cell (i, j) . In our experiments, the salient regions are detected by the algorithm in [24].

The energy function E_{sr} for the right image is defined in a similar way. Thus, the shape preservation term can be formulated as

$$E_s = (E_{s1} + E_{sr})/2. \quad (18)$$

d) Energy optimization: To obtain $\widehat{V_{i,j,l}^2}$ and $\widehat{V_{i,j,r}^2}$, we minimize the total energy including all the above energy terms as

$$E = E_h + E_g + E_y + E_d + E_s. \quad (19)$$

Since the energy E is a quadratic function of $\widehat{V_{i,j,l}^2}$ and $\widehat{V_{i,j,r}^2}$, the minimization problem is solved via a sparse linear solver.

Equations (13) and (14) can be approximated as

$$E_y = \sum_{(\widehat{f_{i,l}^2}, \widehat{f_{i,r}^2}) \in F^3} \left\| \widehat{f_{i,l,y}^2} - \frac{(\widehat{f_{i,l,y}^2} + \widehat{f_{i,r,y}^2})}{2} \right\|^2 + \sum_{(\widehat{f_{i,l}^2}, \widehat{f_{i,r}^2}) \in F^3} \left\| \widehat{f_{i,r,y}^2} - \frac{(\widehat{f_{i,l,y}^2} + \widehat{f_{i,r,y}^2})}{2} \right\|^2 \quad (20)$$

$$E_d = \sum_{(\widehat{f_{i,l}^2}, \widehat{f_{i,r}^2}) \in F^3} \left\| \widehat{f_{i,l,x}^2} - \frac{(\widehat{f_{i,l,x}^2} + \widehat{f_{i,r,x}^2})}{2} - \frac{\varphi(d_i)}{2} \right\|^2 + \sum_{(\widehat{f_{i,l}^2}, \widehat{f_{i,r}^2}) \in F^3} \left\| \widehat{f_{i,r,x}^2} - \frac{(\widehat{f_{i,l,x}^2} + \widehat{f_{i,r,x}^2})}{2} + \frac{\varphi(d_i)}{2} \right\|^2. \quad (21)$$

Then, the coordinates of the grid vertices ($\widehat{V_{i,j,l}^2}$ and $\widehat{V_{i,j,r}^2}$) are uncorrelated and can be solved separately with the same system matrix and different right-hand sides. Once the output grid cells are obtained, the bilinear interpolation algorithm [46] is used to achieve the final result.

C. Image Blending

The last step of the proposed method is image blending. An extended seam-cutting-based method is applied to find a seam in the aligned image, and then, the multiband blending algorithm [17] is deployed to create the final stitching result. To stitch two stereoscopic images together, we need to find a seam in the overlapping region, which is implemented by dynamic programming. Similar to the seam carving method [49], the cumulative energy $M(i, j)$ for connected seams of each entry (i, j) in the first row is computed as

$$M_{i,j} = e(i, j) + \min(M_{i-1,j-1}, M_{i-1,j}, M_{i-1,j+1}) \quad (22)$$

where $e(i, j)$ can be formulated as

$$e(i, j) = \|\widehat{I_{OL}^2}(i, j) - \widehat{I_{OL}^1}(i, j)\| + \|\widehat{G_{OL}^2}(i, j) - \widehat{G_{OL}^1}(i, j)\| \quad (23)$$

where $\widehat{I_{OL}^2}$ and $\widehat{I_{OL}^1}$ denote the overlapping regions of the warped images $\widehat{I_L^2}$ and $\widehat{I_L^1}$, respectively. $\widehat{G_{OL}^2}(i, j)$ and $\widehat{G_{OL}^1}(i, j)$ denote the gradients of overlapping regions of the warped images ($\widehat{I_L^2}$ and $\widehat{I_L^1}$), respectively. At the end of this process, the minimum value of the last row in $M_{i,j}$ indicates the end of the minimal connected vertical seam.

With the above process, the seam in the overlapping region of the left images ($\widehat{I_L^2}$ and $\widehat{I_L^1}$) is found. The seam in the overlapping region of the right images ($\widehat{I_R^2}$ and $\widehat{I_R^1}$) can be found correspondingly.

After the seams in overlapping regions are found, image blending is used to blend the left (right) images. Despite the overlapping region corresponding to the same scene, the gray values in the two images might not be exactly the same due to the differences in imaging conditions. In the seam position of the overlapping region, an obvious splicing gap will be produced. Therefore, image fusion was used to eliminate the splicing gap and make the image stitching results more natural. In this paper, the multiband blending algorithm [17] is employed to get a smooth resultant image.

IV. EXPERIMENTAL RESULTS

In this section, we conduct some comparison experiments to demonstrate the advantages of the proposed method.

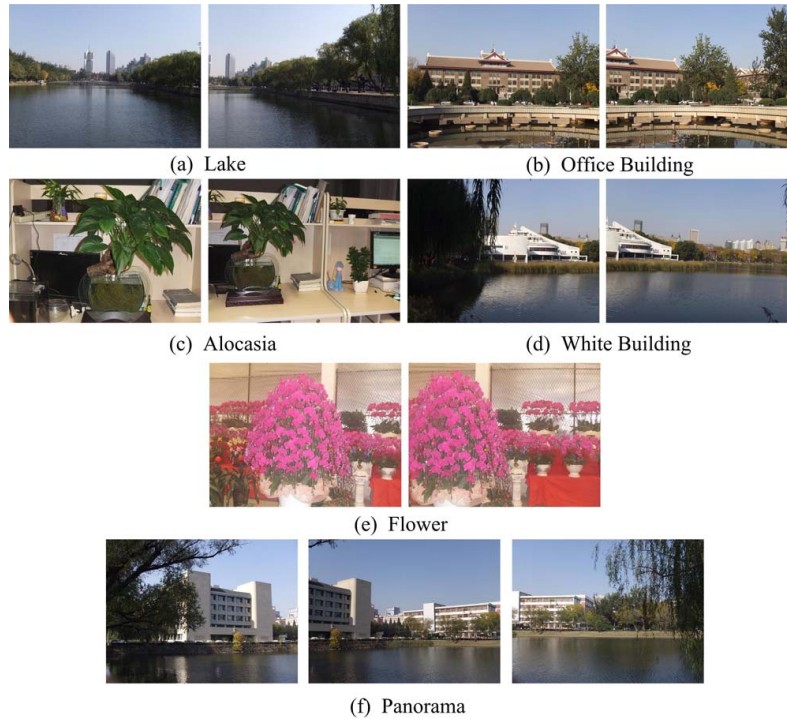


Fig. 4. Six test cases (left images) used in the image stitching.

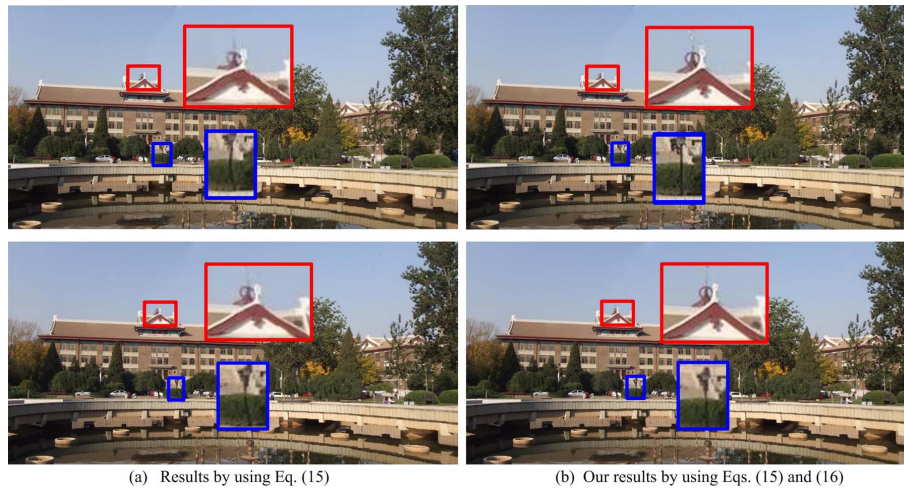


Fig. 5. Comparison of alternative energy terms. First row: stitched left images. Second row: stitched right images.

Since there is no stereoscopic image database available for image stitching, we have constructed a database composed of six stereoscopic images: *Lake*, *Office Building*, *Alocasia*, *White Building*, *Flower*, and *Panorama*. The stereoscopic images were captured using a FUJIFILM REAL 3D W1 digital camera. Fig. 4 shows the six test cases used in our experiments. In Sections IV-A and IV-B, we first report the experimental comparison results, including the comparison of alternative energy terms, the comparison between our method and the three representative image stitching methods: APAP [4], Du's method [41], and Zhang's method [43]. The APAP method is a 2D image stitching method, which better addresses the problem of image alignment compared with the other 2D image stitching methods. Du's method could change

stereoscopic image perspective and apply to stereoscopic image stitching. Zhang's method could directly stitch stereoscopic images. Following the experiment, we show the applications of the proposed stereoscopic image stitching framework. Limited by the medium, the results shown in this section are presented as red/cyan color anaglyph images.

A. Comparisons

In this paper, we use (15) and (16) to reduce distortion and keep the smoothness of warping. To evaluate the effectiveness, we conduct one comparison experiment for the alternative smooth energy term. As shown in Fig. 5, the method using only (15) produces distortion [see the enlarged regions in Fig. 5(a)]. The region on the roof of the building is distorted

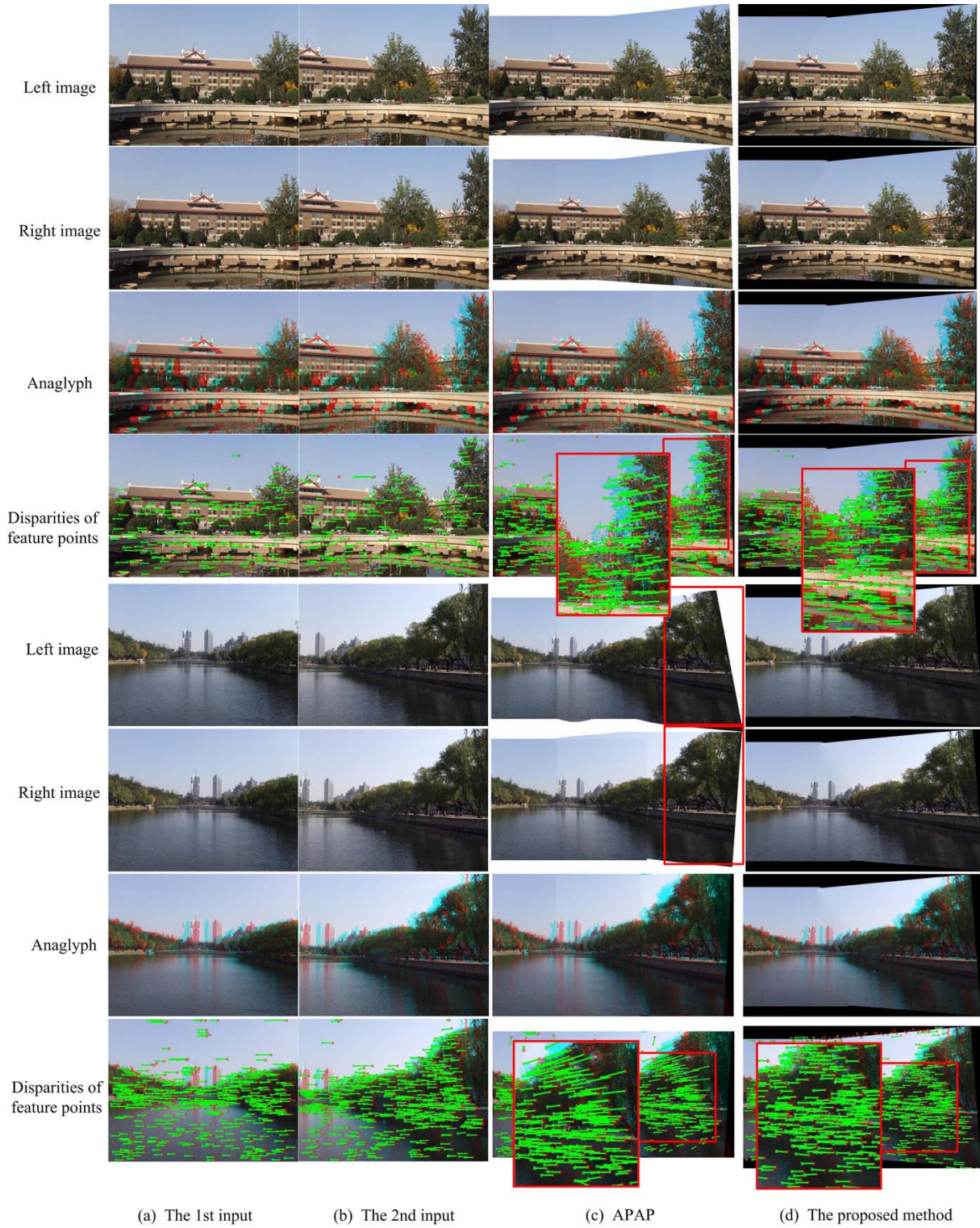


Fig. 6. Comparison results among different stitching approaches. From top to bottom: left image, right image, anaglyph image, and anaglyph image with feature points and disparities.

and the tree is bent. In contrast, the proposed method [Fig. 5(b)], which uses (15) and (16), makes a smoother, more naturalistic stitched image.

We conduct other comparison experiments using the proposed method with other existing ones on a range of challenging stereoscopic image pairs. The proposed method is compared with the native 2D image stitching algorithm of

APAP [4] by applying it independently on the left and right images. The stereoscopic image stitching methods in [41] and [43] are also used in this comparison experiment.

Fig. 6 shows the comparison results of APAP [4] and the proposed method for *Office Building* and *Lake*. The anaglyph stereoscopic images (as shown in the third row in Fig. 6) are obtained by cropping the raw stitching results, and the

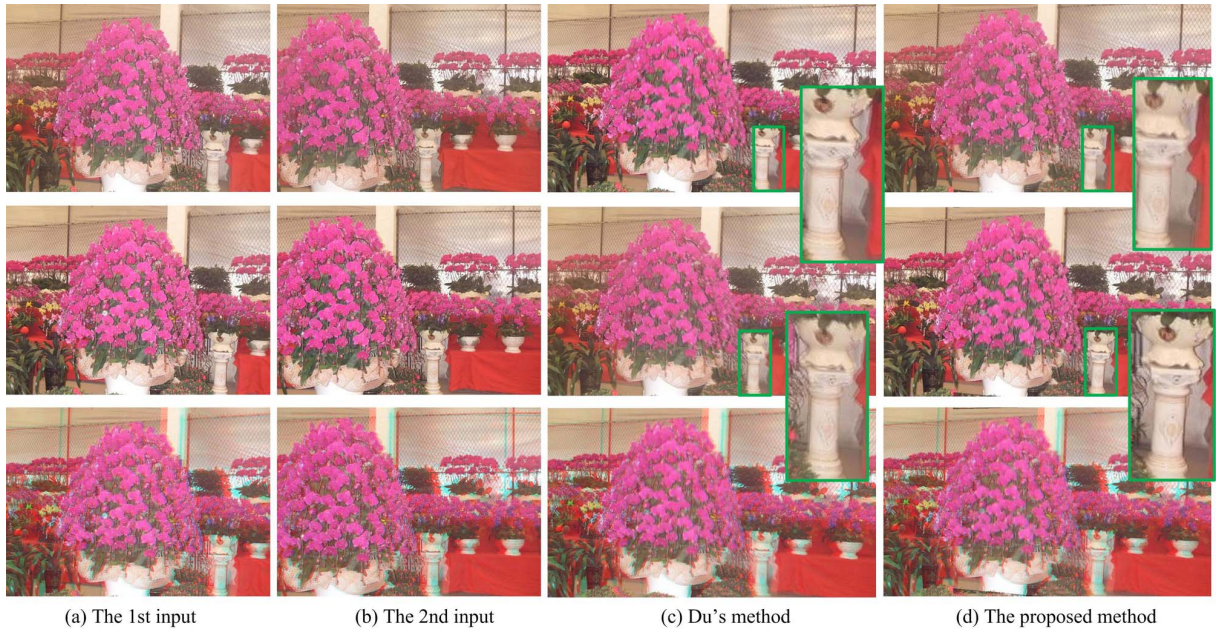


Fig. 7. Comparison results from different methods. From top to bottom: left image, right image, and anaglyph image.

anaglyph images with feature points and disparities (as shown in the fourth row in Fig. 6) show the metric of vertical disparity, where the horizontal green lines denote that the vertical disparity of corresponding feature points is zero. As shown in Fig. 6, the APAP could well align the left and right images separately, while they could not produce a good stereoscopic image due to the lack of stereoscopic constraints. Since APAP stitches the left and right images separately, the difference in the projective transformation is large. The undesired vertical disparity and inconsistent image sizes lead to incorrect and inconsistent 3D perception. As shown in Fig. 6(c), the vertical disparity mainly appears in the right part of the anaglyph image, since APAP changes one of the two images. In contrast to APAP, the proposed method ensures the consistency of image size by a prewarping method based on the uniform homography matrix. In addition, the proposed method guarantees image alignment and reduces vertical disparities in a content-aware manner. The proposed method [Fig. 6(d)] not only successfully aligns the two images, but also obtains a good stereoscopic image with fewer vertical disparities.

Fig. 7 shows the comparison results of Du's method [41] and the proposed method. Du's method can change the perspective of one stereo image by changing the stereo camera orientation, and, moreover, can be directly stitched to construct a stereoscopic panorama. Du's method could lead to a good result with little vertical disparity. However, their method cannot ensure the alignment, since it only considers to change the perspective of one stereoscopic image. In addition, Du's method produces ghosting [see the enlarged regions in Fig. 7(c)]. In contrast to Du's method, the proposed method [Fig. 7(d)], which considers the image alignment term and the stereoscopic constraint term, not only successfully obtains a good stereoscopic image with fewer vertical disparities, but also aligns the two images.

Fig. 8 shows the comparison results of Zhang's method [43] and the proposed method. Zhang's method [Fig. 8(c)] can

align the input left and right images without ghosts. However, it only warps the right images instead of warping both the two images according to the stitched left images and the target disparity map to adjust disparities and keep the consistency, which may produce some distortions. As shown in the second row in Fig. 8(c) (i.e., the stitched right image), there are some distortions in the enlarged region. In contrast, the proposed method [Fig. 8(d)], which uses uniform homography and content-preserving warping method to warp both the images, can effectively reduce the distortion [see the enlarged region in Fig. 8(d)].

Fig. 9 shows the results of the proposed method for another scene. It can be observed from Fig. 9 that the proposed method not only aligns the input images, which is free from the ghosting artifacts, but also reduces vertical disparity, which avoids uncomfortable viewing experience.

The viewing experience of stereoscopic images is affected by various factors [50], and there is no general objective measurement of stereoscopic visual quality. However, it is widely accepted that large vertical disparities increase visual discomfort. Here, we show the average absolute vertical disparity (AVD) for our method and other existing methods. The AVD is computed from the sparse corresponding SIFT features. The AVD is defined as

$$AVD = \frac{\sum_{i=1}^{N_s} |VD_i|}{N_s} \quad (24)$$

where $|VD_i|$ denotes the AVD of the i th feature point and N_s denotes the total number of feature points. Table I shows the AVD of all the six stitched stereoscopic images. From Table I, we can observe that the AVD from APAP is greater than that of the proposed method in all of the image pairs. The main reason is that the APAP does not consider the stereoscopic constraints. While Du's method considers the stereoscopic constraints, the AVD from Du's method is also

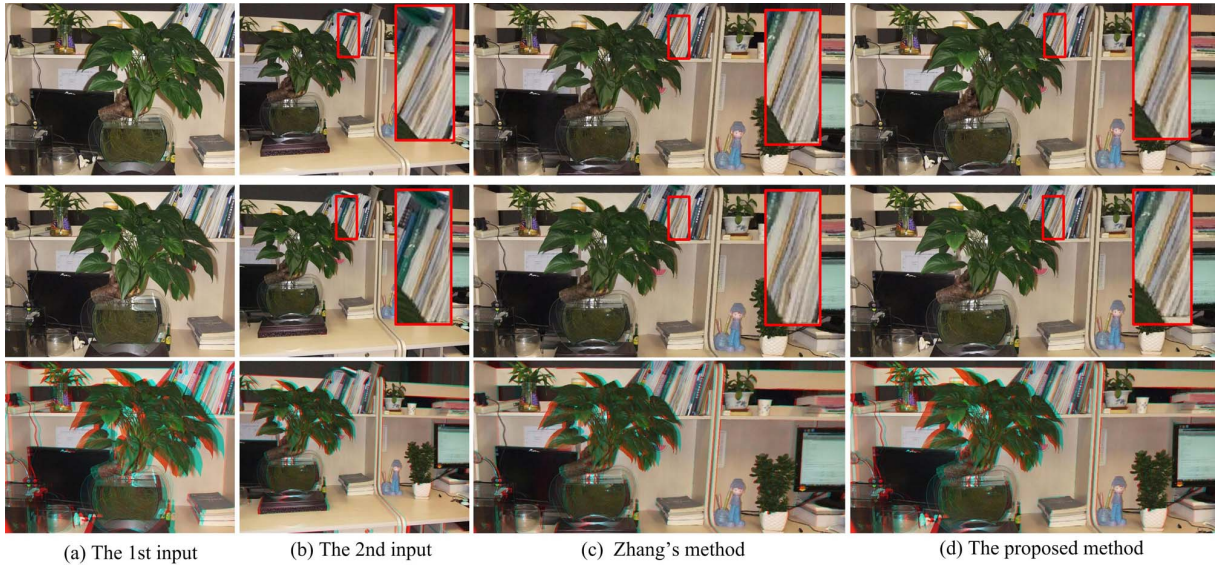


Fig. 8. Comparison results with Zhang's method. From top to bottom: left image, right image, and anaglyph image.



Fig. 9. Stitched results of the proposed method for another scene. From left to right are the first input stereoscopic image pairs, the second input stereoscopic image pairs, and the stitching results of the proposed method, while from top to bottom are the left image, the right image, and the anaglyph image.

TABLE I
AVERAGE VERTICAL DISPARITY (/PIXEL)

Dataset	APAP	Du's method	Zhang's method	Proposed
Lake	2.69	0.95	0.92	0.90
Office Building	1.17	1.63	1.09	1.03
Alocasia	6.37	1.31	1.26	1.21
White Building	2.80	1.79	1.35	1.36
Flower	2.56	1.30	1.13	1.08
Panorama	2.81	1.43	1.40	1.39

larger than that of the proposed method. The main reason is that Du's method cannot guarantee the disparity consistency in the overlapping regions of the images. Zhang's method also produces larger vertical disparities than our method. The proposed method, which uses uniform homography and content-preserving warping to warp both the images, reduces vertical disparity by considering the stereoscopic constraints.

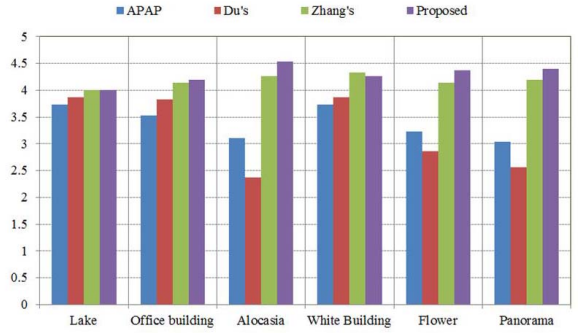


Fig. 10. Average evaluation score.

To better demonstrate the advantages of the proposed method, we also conducted a user study [51], [52] to compare the visual quality of the stereoscopic images obtained by APAP, Du's method, Zhang's method, and the proposed method. A total of 15 participants with normal stereoscopic vision were invited to participate in the user study. These participants were asked if they felt comfortable with the factors, including ghost, distortion, and eye fatigue. They were asked to rate from 1 (very uncomfortable) to 5 (very comfortable). Table II shows the subjective assessment standard of stitched stereoscopic images.

We compute the average scores and show them in Fig. 10. From Fig. 10, we can observe that the proposed method outperforms the three other methods. From these experimental results, we can observe that the proposed method provides more comfortable viewing effects for the stitched results compared with other existing ones by considering image alignment and the stereoscopic constraints.

B. Applications

In this section, we apply the proposed method in various applications of stereoscopic image processing to demonstrate the advantages of the proposed method. We first present the

TABLE II
SUBJECTIVE ASSESSMENT STANDARD OF STITCHED
STEREOSCOPIC IMAGES

Rank	Assessment Standard	Assessment Score
Excellent	Very easy to perceive stereoscopic sense and without any distortion and ghost; stereoscopic sense is comfortable, natural, and in accordance with visual experience.	5
Good	Easy to perceive stereoscopic sense and without any distortion and ghost; slight influence to the comfort of stereoscopic sense; stereoscopic sense is in accordance with visual experience.	4
Fair	Slightly feeling of ghost and distortion; stereoscopic sense is basically in accordance with visual experience.	3
Poor	Mostly feeling of ghost and distortion; feel uncomfortable with the stereoscopic sense.	2
Bad	Intense feeling of ghost and distortion; feel annoyed with the stereoscopic sense.	1

production scenarios of disparity adjustment. Then, we illustrate the extension of the proposed method in the application of full panoramas.

1) *Disparity Adjustment*: The disparity range of the pre-aligned stereoscopic images may be not the optimal. In order to provide a strong visual impact, large negative disparities need to be set to particular scenes. On the contrary, stereoscopic videos with a period of large negative disparities may cause visual fatigue. Therefore, the disparity has to be adjusted. The proposed method can allow for the adjustment of disparity. As mentioned in Section III-B, since $\varphi(d_i)$ in (14) can be adjusted, the proposed method can also be used to change the perceived depth of a stitching stereoscopic image by changing the disparity mapping $\varphi(d_i)$, similar with the technique in [39]. Our method can compute a deformation of the prealigned stereoscopic images $\overline{I}_L^2$ and $\overline{I}_R^2$ so as to meet the target disparities. Fig. 11 shows an example of disparity adjustment. Here, $\varphi(d_i) = 0.25(f_{i,l,x}^2 - f_{i,r,x}^2)$. It can be observed from the result that the proposed method can adjust the disparity as expected.

2) *Constructing Stereoscopic Image With a Larger Horizon*: With the popularity of stereoscopic images, the techniques of producing and editing panoramic stereoscopic images would be widely used. A panorama stereoscopic image is a wide-angle view or representation of a physical space. The proposed method can construct a stereoscopic image with a larger horizon.

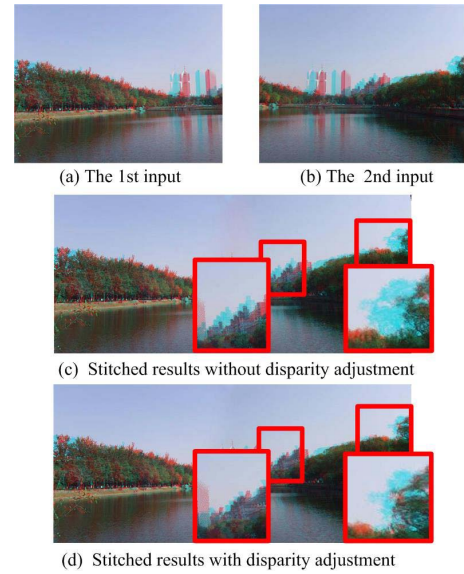


Fig. 11. Disparity adjustment.

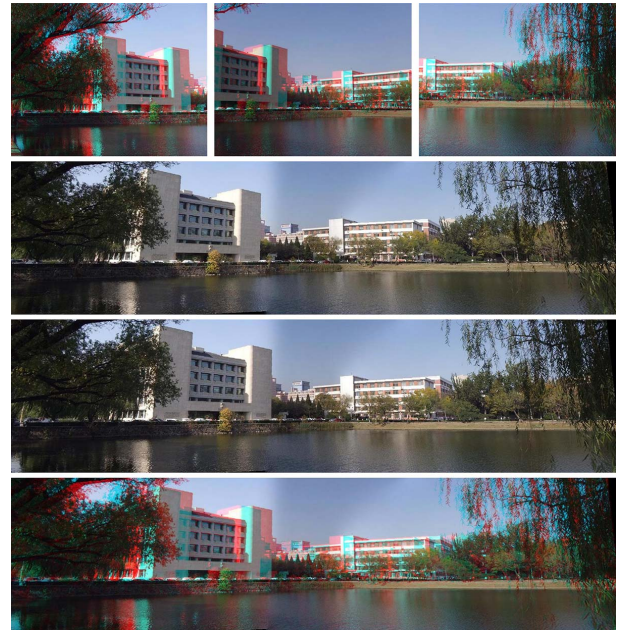


Fig. 12. Stereoscopic panorama image. From top to bottom: input stereoscopic images, stitched left image, stitched right image, and the anaglyph image.

To obtain a stereoscopic image with a larger horizon, we can stitch a set of overlapping images incrementally using the proposed method. Given more than two stereoscopic images, we first choose a central image to initialize the stitched image. Then, the other images are incrementally stitched via the proposed method. Fig. 12 shows an example of the panorama stereoscopic image. This achieves a good panorama stereoscopic image without ghosts. In addition, to acquire 360° stereoscopic panoramas, a spherical or cylindrical parametrization model for stereoscopic images may be introduced to prealign multiple images in the future work.

V. CONCLUSION

In this paper, we have presented a novel hybrid warping model for stereoscopic image stitching. To ensure the

consistent changes, a uniform homography is proposed to prewarp the left and right images. The content-preserving warping is introduced to locally refine the alignment and reduce the vertical disparity. Finally, a seam-cutting-based algorithm is designed to find a blending seam, and the multi-band blending algorithm is used to produce the final stitched image. Experimental results on stereoscopic image stitching have shown the effectiveness of the proposed method and its applications.

ACKNOWLEDGMENT

The authors would like to thank Dr. S. A. Cholewiak for help with the English language. They would also like to thank the anonymous reviewers for their valuable comments that significantly helped in improving the presentation of this paper.

REFERENCES

- [1] H.-Y. Shum and R. Szeliski, "Construction and refinement of panoramic mosaics with global and local alignment," in *Proc. 6th Int. Conf. Comput. Vis.*, 1998, pp. 953–958.
- [2] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *Int. J. Comput. Vis.*, vol. 74, no. 1, pp. 59–73, 2007.
- [3] R. Szeliski, "Image alignment and stitching: A tutorial," *Found. Trends Comput. Graph. Vis.*, vol. 2, no. 1, pp. 1–104, 2006.
- [4] J. Zaragoza, T.-J. Chin, Q.-H. Tran, M. S. Brown, and D. Suter, "As-projective-as-possible image stitching with moving DLT," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 7, pp. 1285–1298, Jul. 2014.
- [5] W.-Y. Lin, S. Liu, Y. Matsushita, T.-T. Ng, and L.-F. Cheong, "Smoothly varying affine stitching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 345–352.
- [6] 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.
- [7] H.-C. Huang and Y.-P. Hung, "Panoramic stereo imaging system with automatic disparity warping and seaming," *Graph. Models Image Process.*, vol. 60, no. 3, pp. 196–208, 1998.
- [8] H. Tang and Z. Zhu, "Content-based 3-D mosaics for representing videos of dynamic urban scenes," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 2, pp. 295–308, Feb. 2012.
- [9] C. Richardt, Y. Pritch, H. Zimmer, and A. Sorkine-Hornung, "Megastereo: Constructing high-resolution stereo panoramas," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1256–1263.
- [10] Z. Zhu, A. R. Hanson, and E. M. Riseman, "Generalized parallel-perspective stereo mosaics from airborne video," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 226–237, Feb. 2004.
- [11] L. E. Gurrieri and E. Dubois, "Stereoscopic cameras for the real-time acquisition of panoramic 3D images and videos," *Proc. SPIE*, vol. 8648, pp. 86481W-1–86481W-17, Mar. 2013.
- [12] S. S. Hwang, W.-J. Kim, J. Yoo, and S. D. Kim, "Visual hull-based geometric data compression of a 3-D object," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 7, pp. 1151–1160, Jul. 2015.
- [13] Y. Zhao, C. Zhu, Z. Chen, D. Tian, and L. Yu, "Boundary artifact reduction in view synthesis of 3D video: From perspective of texture-depth alignment," *IEEE Trans. Broadcast.*, vol. 57, no. 2, pp. 510–522, Jun. 2011.
- [14] J. Lei, S. Li, C. Zhu, M.-T. Sun, and C. Hou, "Depth coding based on depth-texture motion and structure similarities," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 2, pp. 275–286, Feb. 2015.
- [15] A. Agarwala *et al.*, "Interactive digital photomontage," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 294–302, 2004.
- [16] A. Eden, M. Uyttendaele, and R. Szeliski, "Seamless image stitching of scenes with large motions and exposure differences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2006, pp. 2498–2505.
- [17] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Trans. Graph.*, vol. 2, no. 4, pp. 217–236, 1983.
- [18] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 313–318, Jul. 2003.
- [19] R. Szeliski and H.-Y. Shum, "Creating full view panoramic image mosaics and environment maps," in *Proc. 24th SIGGRAPH*, 1997, pp. 251–258.
- [20] R. Carroll, M. Agrawala, and A. Agarwala, "Optimizing content-preserving projections for wide-angle images," *ACM Trans. Graph.*, vol. 28, no. 3, 2009, Art. no. 43.
- [21] J. Wei, C.-F. Li, S.-M. Hu, R. R. Martin, and C.-L. Tai, "Fisheye video correction," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 10, pp. 1771–1783, Oct. 2012.
- [22] R. Carroll, A. Agarwala, and M. Agrawala, "Image warps for artistic perspective manipulation," *ACM Trans. Graph.*, vol. 29, no. 4, 2010, Art. no. 127.
- [23] Y. Fang, J. Wang, M. Narwaria, P. Le Callet, and W. Lin, "Saliency detection for stereoscopic images," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2625–2636, Jun. 2014.
- [24] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, pp. 1915–1926, Oct. 2012.
- [25] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in *Advances in Neural Information Processing Systems*. Cambridge, MA, USA: MIT Press, 2006, pp. 545–552.
- [26] Z. Liu, W. Zou, and O. Le Meur, "Saliency tree: A novel saliency detection framework," *IEEE Trans. Image Process.*, vol. 23, no. 5, pp. 1937–1952, May 2013.
- [27] R. Song, Y. Liu, R. R. Martin, and P. L. Rosin, "Mesh saliency via spectral processing," *ACM Trans. Graph.*, vol. 33, no. 1, 2014, Art. no. 6.
- [28] Y.-S. Wang, C.-L. Tai, O. Sorkine, and T.-Y. Lee, "Optimized scale-and-stretch for image resizing," *ACM Trans. Graph.*, vol. 27, no. 5, 2008, Art. no. 118.
- [29] M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir, "A comparative study of image retargeting," *ACM Trans. Graph.*, vol. 29, no. 6, 2010, Art. no. 160.
- [30] Y. Fang, W. Lin, Z. Chen, C.-M. Tsai, and C.-W. Lin, "A video saliency detection model in compressed domain," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 1, pp. 27–38, Jan. 2014.
- [31] C. Guo and L. Zhang, "A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression," *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 185–198, Jan. 2010.
- [32] W. Hu, R. Hu, N. Xie, H. Ling, and S. Maybank, "Image classification using multiscale information fusion based on saliency driven nonlinear diffusion filtering," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1513–1526, Apr. 2014.
- [33] C.-H. Chang, C.-K. Liang, and Y.-Y. Chuang, "Content-aware display adaptation and interactive editing for stereoscopic images," *IEEE Trans. Multimedia*, vol. 13, no. 4, pp. 589–601, Aug. 2011.
- [34] J. W. Yoo, S. Yea, and I. K. Park, "Content-driven retargeting of stereoscopic images," *IEEE Signal Process. Lett.*, vol. 20, no. 5, pp. 519–522, May 2013.
- [35] F. Liu, M. Gleicher, H. Jin, and A. Agarwala, "Content-preserving warps for 3D video stabilization," *ACM Trans. Graph.*, vol. 28, no. 3, 2009, Art. no. 44.
- [36] S.-J. Luo, I.-C. Shen, B.-Y. Chen, W.-H. Cheng, and Y.-Y. Chuang, "Perspective-aware warping for seamless stereoscopic image cloning," *ACM Trans. Graph.*, vol. 31, no. 6, 2012, Art. no. 182.
- [37] F. Liu, Y. Niu, and H. Jin, "Casual stereoscopic photo authoring," *IEEE Trans. Multimedia*, vol. 15, no. 1, pp. 129–140, Jan. 2013.
- [38] R. F. Tong, Y. Zhang, and K. L. Cheng, "StereoPasting: Interactive composition in stereoscopic images," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 8, pp. 1375–1385, Aug. 2013.
- [39] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and M. Gross, "Nonlinear disparity mapping for stereoscopic 3D," *ACM Trans. Graph.*, vol. 29, no. 4, 2010, Art. no. 75.
- [40] T. Yan, R. W. Lau, Y. Xu, and L. Huang, "Depth mapping for stereoscopic videos," *Int. J. Comput. Vis.*, vol. 102, nos. 1–3, pp. 293–307, 2013.
- [41] S.-P. Du, S.-M. Hu, and R. R. Martin, "Changing perspective in stereoscopic images," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 8, pp. 1288–1297, Aug. 2013.
- [42] Y. Niu, W.-C. Feng, and F. Liu, "Enabling warping on stereoscopic images," *ACM Trans. Graph.*, vol. 31, no. 6, 2012, Art. no. 183.
- [43] F. Zhang and F. Liu, "Casual stereoscopic panorama stitching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 2002–2010.
- [44] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [45] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
- [46] P. S. Heckbert, "Fundamentals of texture mapping and image warping," M.S. thesis, Dept. Elect. Eng. Comput. Sci., Univ. California, Berkeley, Berkeley, CA, USA, 1989.

- [47] Z. Gu, W. Lin, B.-S. Lee, and C. T. Lau, "Rotated orthogonal transform (ROT) for motion-compensation residual coding," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4770–4781, Dec. 2012.
- [48] Z. Gu, W. Lin, B.-S. Lee, and C. Lau, "Low-complexity video coding based on two-dimensional singular value decomposition," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 674–687, Feb. 2012.
- [49] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," *ACM Trans. Graph.*, vol. 26, no. 3, 2007, Art. no. 10.
- [50] M. Lambooi, M. Fortuin, I. Heynderickx, and W. IJsselstein, "Visual discomfort and visual fatigue of stereoscopic displays: A review," *J. Imag. Sci. Technol.*, vol. 53, no. 3, pp. 30201–1–30201–14, 2009.
- [51] *Methodology for the Subjective Assessment of the Quality of Television Pictures*, document Rec. ITU-R BT.500-11, International Telecommunication Union, Geneva, Switzerland, 2002.
- [52] J. Lei, S. Li, B. Wang, K. Fan, and C. Hou, "Stereoscopic visual attention guided disparity control for multiview images," *J. Display Technol.*, vol. 10, no. 5, pp. 373–379, May 2014.



Weiqing Yan received the B.S. degree from the School of Mathematics and Statistics Science, Ludong University, Yantai, China, in 2010. She is currently working toward the Ph.D. degree with the School of Electronic Information Engineering, Tianjin University, Tianjin, China.

She has been a Visiting Student with University of California at Berkeley, Berkeley, CA, USA, since 2015. Her research interests include 3D image editing, computer graphic, and computer vision.



Chunping Hou received the M.Eng. and Ph.D. degrees in electronics engineering from Tianjin University, Tianjin, China, in 1986 and 1998, respectively.

She has been a Faculty Member with the School of Electronic and Information Engineering, Tianjin University, since 1986, where she is currently a Full Professor and the Director of the Broadband Wireless Communications and 3D Imaging Institute. Her research interests include 3D image processing, 3D display, wireless communication, and the design and applications of communication systems.



Jianjun Lei (M'11) received the Ph.D. degree in signal and information processing from Beijing University of Posts and Telecommunications, Beijing, China, in 2007.

He was a Visiting Researcher with the Department of Electrical Engineering, University of Washington, Seattle, WA, USA, from 2012 to 2013. He is currently a Professor with the School of Electronic Information Engineering, Tianjin University, Tianjin, China. His research interests include 3D video processing, 3D display, and computer vision.

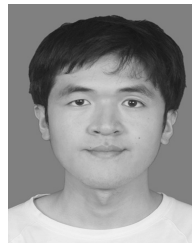


Yuming Fang (M'13) received the B.E. degree from Sichuan University, Chengdu, China; the M.S. degree from Beijing University of Technology, Beijing, China; and the Ph.D. degree in computer engineering from Nanyang Technological University, Singapore.

He was a Visiting Post-Doctoral Research Fellow with the IRCCyN Laboratory, PolyTech Nantes, Nantes, France; University of Nantes, Nantes; University of Waterloo, Waterloo, ON, Canada; and Nanyang Technological University. He is currently

a Faculty Member with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, China. His research interests include visual attention modeling, visual quality assessment, image retargeting, computer vision, and 3D image/video processing.

Dr. Fang was a Secretary of the Ninth Joint Conference on Harmonious Human Machine Environment. He was also a Special Session Organizer of the IEEE International Conference on Visual Communications and Image Processing 2013 and Quality of Multimedia Experience 2014.



Zhouye Gu (M'14) received the B.Sc. degree in electronic information science and technology from Nanjing University, Nanjing, China, and the Ph.D. degree in computer engineering from Nanyang Technological University, Singapore.

He was a Visiting Student with University of Washington, Seattle, WA, USA, in 2012 and a Research Fellow with Santa Clara University, Santa Clara, CA, USA, from 2012 to 2014. He is currently a Staff Advanced Research Engineer with ARRIS Group Inc., San Diego, CA, USA. His

research interests include video coding and processing, machine learning, and data mining.

Dr. Gu was the Technical Program Co-Chair of Ubi-Media Computing 2014 and the Area Chair of the Institute for Computational and Mathematical Engineering 2013.



Nam Ling (S'88–M'90–SM'99–F'08) received the B.Eng. degree from National University of Singapore, Singapore, in 1981, and the M.S. and Ph.D. degrees from University of Louisiana, Lafayette, LA, USA, in 1985 and 1989, respectively.

He was an Associate Dean of the School of Engineering with Santa Clara University, Santa Clara, CA, USA, from 2002 to 2010. He is currently the Sanfilippo Family Chair Professor and the Chair of the Department of Computer Engineering with Santa Clara University. He is also a Consulting

Professor with the National University of Singapore; a Guest Professor with Tianjin University, Tianjin, China, and Shanghai Jiao Tong University, Shanghai, China; a Cuiying Chair Professor with Lanzhou University, Gansu, China; and a Distinguished Professor with Xi'an University of Posts and Telecommunications (XUPT), Beijing, China. He has authored or co-authored over 180 publications and standard contributions, including two books in the fields of video coding and systolic arrays. He holds over 15 U.S. patents.

Dr. Ling is an Institution of Engineering and Technology (IET) Fellow. He received the IEEE International Conference on Consumer Electronics Best Paper Award (First Place). He was a recipient of six awards from Santa Clara University, four at the University level (Outstanding Achievement, Recent Achievement in Scholarship, President's Recognition, and Sustained Excellence in Scholarship), and two at the School/College level (Researcher of the Year and Teaching Excellence). He was named an IEEE Distinguished Lecturer (twice) and was also an Asia Pacific Signal and Information Processing Association (APSIPA) Distinguished Lecturer. He was a Keynote Speaker of the IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), the International Workshop on Video Coding and Video Processing (VCVP) (twice), the Joint Conferences on Pervasive Computing, the IEEE International Conference on Adaptive Science & Technology, the IEEE Conference on Industrial Electronics and Applications (ICIEA), the IET Conference on Frontier Computing Ubi-Media Computing (Umedia), the IEEE Umedia, and the Workshop on Image and Video Processing for Applications in Electronic Information Scene Investigation at XUPT, and a Distinguished Speaker of the IEEE ICIEA. He has served as a General Chair/Co-Chair of the IEEE Symposium on High Performance Chips, VCVP (twice), the IEEE International Conference on Multimedia and Expo, the IEEE Umedia (three times), and the IEEE International Workshop on Signal Processing Systems (SiPS). He has served as a Technical Program Co-Chair of the IEEE International Symposium on Circuits and Systems, the APSIPA Annual Summit and Conference, the IEEE APCCAS, the IEEE SiPS (twice), Digital and Computational Video, and the IEEE International Conference on Visual Communications and Image Processing. He was a Technical Committee Chair of the IEEE Circuits and Systems for Communications Technical Committee and the IEEE Computer Society Technical Committee on Microprocessors and Microcomputers, and has also served as a Guest Editor or an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, IEEE *Journal of Selected Topics in Signal Processing*, *Journal of Signal Processing Systems* (Springer), and *Multidimensional Systems and Signal Processing*.