IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 6, 2018, accepted January 6, 2019, date of publication January 11, 2019, date of current version March 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2892098

User Controllable Content Retargeting and Depth Adaptation for Stereoscopic Display

FENG SHAO[1], (Member, IEEE), LIBO SHEN[1], QIUPING JIANG[1], (Student Member, IEEE), FUCUI LI[1], AND YO-SUNG HO[2], (Fellow, IEEE)

[1]Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China
[2]School of Information and Communications, Gwangju Institute of Science and Technology, Gwangju 500-712, South Korea

Corresponding author: Feng Shao (shaofeng@nbu.edu.cn)

This work was supported in part by the Natural Science Foundation of China under Grant 61622109, in part by the Zhejiang Natural Science Foundation of China under Grant R18F010008, in part by the Natural Science Foundation of Ningbo under Grant 2017A610112, and in part by the K. C. Wong Magna Fund in Ningbo University.

ABSTRACT This paper presents a content retargeting and depth adaptation method for the stereoscopic image. Our method allows the user to specify the retargeting scenarios to enhance the user's visual experience. Toward this end, we proposed a warping framework that takes content retargeting, depth adaptation, and interactive editing into account simultaneously for stereoscopic images. From the viewpoint of accuracy, comfortability, and controllability, we exploit the complementary relationship among image quality, depth quality, and important content and propose a grid optimization framework to fuse the three indictors. The experimental results demonstrate that our method achieves a preferable tradeoff among accuracy, comfortability, and controllability of information presentation in retargeting, obtaining satisfactory visual experience for users.

INDEX TERMS Stereoscopic display, content retargeting, depth adaptation, viewing experience.

I. INTRODUCTION

With the rapid development of stereoscopic displays from the television to the mobile devices, especially with the booming virtual reality (VR) researches and applications in the recent years, stereoscopic image/video editing is becoming important in providing more natural viewing experiences for the users on different displays [1]–[3]. However, stereoscopic images with fixed resolutions are unable to provide fully natural visual cues and comfortable viewing experiences on different display resolutions. Stereoscopic image retargeting is able to satisfy the requirements in offering comfortable 3D viewing experience.

2D image retargeting has drawn wide attentions over the recent years, and many methods have been developed, which can be broadly classified into two categories: discrete methods [4]–[6] and continuous methods [7]–[9]. Compared with 2D case, 3D image retargeting is a challenging issue, which not only delivers image quality, but also the impressive 3D depth perception. Therefore, directly applying 2D resizing methods on the left and right images is a straightforward but not a feasible solution. These years have witnessed great efforts in stereoscopic image retargeting, including cropping [10]–[12], seam carving [13]–[15], warping based [16]–[18] methods. However, disparity/depth preservation constraint is utilized in these stereoscopic image retargeting methods, deviating from the motivation of image retargeting to adapt on different display devices. Therefore, only resizing the spatial resolution with depth preservation constraint is not sufficient to account for user's visual experience.

To tackle these challenges, two special requirements for stereoscopic retargeting, i.e., geometric deformation and depth adaption, are introduced to maximize the user's experience. Also, it is critical to demonstrate important content at a proper scale in the resized image. Therefore, accuracy, comfortability and controllability are three important indicators considered in this work to balance the factors affecting 3D visual experience. Here, accuracy is to minimize the geometric deformation, comfortability is to ensure the proper depth perception, and controllability is to allow the user to distribute the content in a proper scale and convey important information freely. As illustrated by the example in Fig.1, our method applies image quality energy to retarget the stereoscopic image, thus avoiding structure degradation (Fig.1(a)),

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

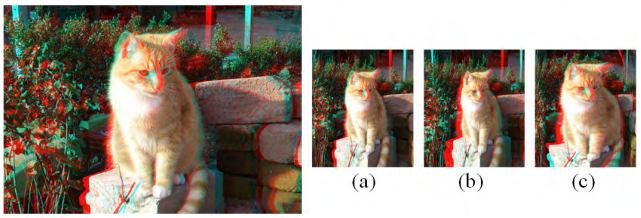


FIGURE 1. Image retargeting generated by different indictors in our method.

and uses stereoscopic image quality energy to adjust the depth range (Fig.1(b)). Our method also applies a cropping window to maintain a balance between the important content and overall completeness in generating the retargeted content (Fig.1(c)). Therefore, from the perspective of information comprehensiveness (also related to the user's visual experience), combination of various indicators can help to integrate the advantages of each component and compensate the drawbacks with each other.

In this paper, we use user controllable retargeting scenarios to enhance user's visual experience in stereoscopic 3D display. To match such requirement, the controllable retargeting scenarios include depth range, viewing distance and cropping window. We obtain content retargeting and depth adaptation according to three retargeting operators: image quality, depth quality, and important content. Firstly, image quality warping operation distributes deformation across an image non-homogeneously according to the visual importance of the grids. Secondly, depth quality warping operation helps to adapt the depth range as well as the projected shape of the grids. Moreover, important content warping operation is to display important content at a proper size and depth range, which aims at providing more space for user to edit important information. The three operators are fused as total energy constraints to optimize the grid deformation and depth adaption process. By this way, the proposed optimization procedure targets to balance the various factors affecting 3D visual experience. The main contributions are summarized as follows:

1) We propose an optimization framework for stereoscopic image retargeting based on energy constraints of image quality, depth quality, and important content, which provide a valuable tool for content retargeting, depth adaptation and interactive editing of stereoscopic images simultaneously.

2) We integrate various user controllable retargeting scenarios, including allowable depth range, viewing distance and cropping window, into the stereoscopic image retargeting framework. The grid coordinates and the perceived depths are simultaneously optimized to address the retargeting scenarios. There is very little previous work using these scenarios.

3) We present a novel system for interactive editing, in which user can edit the size and depth range for the important content within the cropping window. To ensure the completeness of the retargeted content, the system is optimized with other energy constraints in a unified optimization framework.

In the remainder of this paper, we first review the related work in Section II, detail our method in Section III, and finally present results in Section IV and discussion in Section V.

II. RELATED WORK

As discussed, 2D stereoscopic image retargeting can be classified into discrete methods and continuous methods. Overview of state-of-the-art 2D image retargeting methods can refer to [19]. With the same categories for stereoscopic image retargeting, cropping [10]–[12] and seam carving [13]–[15], [20]–[22] are well-known discrete approaches. In cropping, a cropping window or critical region is determined to preserve the aesthetic value [10], or maximize the visually important content [11], [12]. The advantage of cropping methods is that the content outside the cropping window are weakened or even removed while highlighting the clipped content. In seam carving, continuous or discontinuous seams are iteratively carved or inserted to reach the desired size. As typical works in stereoscopic image seam carving, Utsugi *et al.* [13] generated different corresponding seams to maintain consistency and used occluded seams to change the consistency. Basha *et al.* [14] optimized seam caving by taking the visibility relations between pixels in the image pair into account, and obtained geometrically consistent results. Lei *et al.* [15] used pixel fusion technique to adaptively retarget stereo images with flexible aspect ratios while preserving the depth. Besides, different 3D saliency models were proposed to guide seam carving [20]–[22]. However, although this technique has high flexibility in removing pixels, it will cause visual distortion in/between the left and right images.

Compared with the above methods that discretely remove seams or crop borders of an image, continuous methods optimize warping using several deformation and other constraints. Chang *et al.* [16] used single-layer warping-based retargeting algorithm to interactive stereoscopic image editing. The method retains the disparity of the sparse feature points with disparity consistency constraint. However, due to the lack of object deformation and accurate stereo correspondences, the shapes and structures may be distorted after retargeting. Lee *et al.* [17] proposed a multiple-layer warping-based retargeting approach. Their method warps each layer by its own mesh deformation and composites all layers together to form the resized images. The method suffers from the issues of accurate object extraction and object occlusion. Li *et al.* [18] presented a depth-preserving warping-based stereoscopic image retargeting framework that simultaneously preserves the shape of salient objects and the depth of 3D scenes. Lin *et al.* [23] utilized the matched objects between the left and right images to generate an object-based significance map and preserve object consistency. Other relevant works can be found in [24]–[28]. However, most of these methods over-preserved the disparity/depth information.

Recently, besides the above stereoscopic image retargeting methods, many efforts have been made in stereoscopic disparity/depth editing, such as perspective manipulation [29],

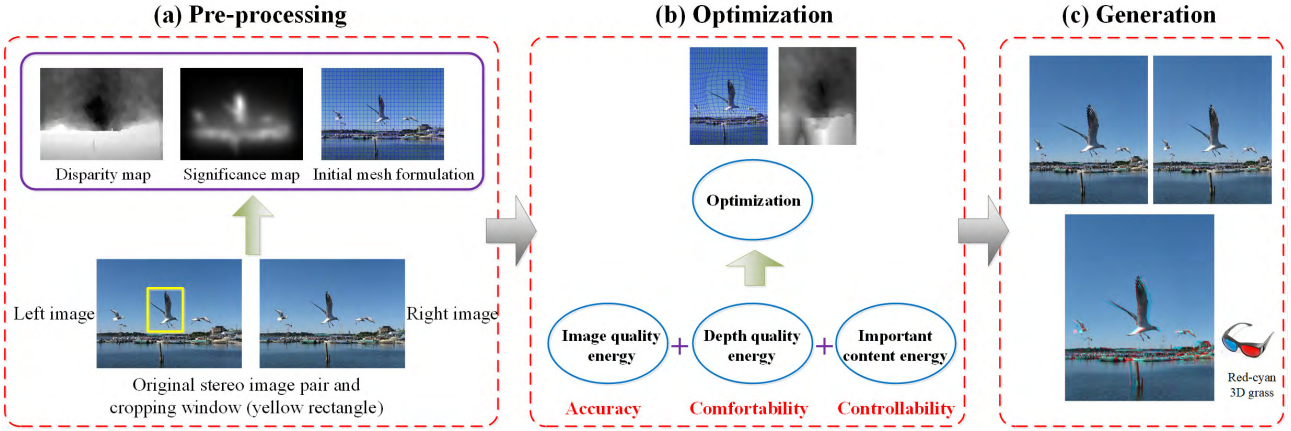


FIGURE 2. Overview of our user controllable stereoscopic image retargeting approach.

stereoscopic composition [3], and disparity mapping [30]. Focused on editing disparity information, Lang *et al*. [30] proposed a series of disparity mapping operators for stereoscopic images and videos to change the disparity ranges. Yan *et al*. [31] proposed a linear depth mapping method through warping to adjust the depth range of a stereoscopic video according to the viewing configuration. Wang *et al*. [32] proposed a mapping optimization method to adjust the disparity to minimize discomfort also by warping-based manipulation. Yan *et al*. [33] proposed a content-aware stereoscopic mesh warping model to determines a scaling factor of salient region by disparity scaling factor. Part *et al*. [34] adjusted the 3D depth of an object via virtual fronto-parallel planar projection in the 3D space. Besides, disparity remapping functions are optimized to control the disparity range [35], [36].

In this work, motivated by the work [16] that simultaneously considers image retargeting and depth adaptation, and the work [37] that incorporates cropping energy into the optimization, we further integrate the cues from image quality, stereoscopic quality and important content to achieve content retargeting, depth adaptation and interactive editing simultaneously. Since Chang's work still prefers retaining the disparity of the corresponding sparse feature points, the perceived depth outside the feature regions are more likely to be distorted, while our method can well eliminate the problem by multi-cue fusion to maximize the user's experience. Essentially, our method belongs to stereoscopic image editing with the ultimate goal to provide an efficient editor for 3D media.

III. USER CONTROLLABLE STEREOSCOPIC IMAGE RETARGETING

Fig. 2 shows the framework of the stereoscopic image retargeting approach, which mainly consists of three functional modules: pre-processing module, warping module and image generation module. The first module is used to generate disparity map and significance map to determine which regions are informative and responsible for interactive selections of cropping window, scaling factors and viewing distance. The second module is to optimize the grid deformation and depth adaption based on energy constraints of image quality, stereoscopic quality, and important content. Minimizing the energy function results in new coordinates and depths of grid vertices, and consequently the retargeted left image is rendered by mapping original image into a new grid structure. At the last module, both left and right images will be resized to its target scale. Different with state-of-the-art stereoscopic image retargeting methods that preserve the original disparity/depth information after retargeting, our method optimizes grid's coordinates and depth ranges simultaneously to match the target displays and user's interaction requirements.

A. PROBLEM FORMULATION

Unlike retargeting in 2D images which only consider the spatial coordinates, stereoscopic retargeting is nontrival due to the added depth perception. Let $\mathbf{V} = \{\mathbf{V}_k, k = 1, \cdots, N\}$ be the vertex set of the grid mesh, in which each mesh $\mathbf{V}_k = \{\mathbf{v}_k^1, \mathbf{v}_k^2, \mathbf{v}_k^3, \mathbf{v}_k^4\}$ is created from four vertexes $\mathbf{v}_k^1$, $\mathbf{v}_k^2$, $\mathbf{v}_k^3$ and $\mathbf{v}_k^4$ (N is the number of meshes in the left image). The vertex set is also defined as $\mathbf{V} = \{\mathbf{v}_i, i = 1, \cdots, M\}$ (M is the number of vertexes of all meshes). The edge set of all vertexes is described as $\mathbf{E} = \{\mathbf{e}_{i,j}\}$. The depth set of all vertexes is defined as $\mathbf{Z} = \{Z_i, i = 1, \cdots, M\}$, in which Z_i denotes the depth value of each vertex. As discussed below, we calculate the perceived depth instead of disparity to incorporate the specific viewing configuration because the user perceives the depth on the 3D space instead of on the image plane. The goal of stereoscopic image retargeting motivated in this work is to change the dimensions of a stereoscopic image pair with size $w \times h$ to a desired size $\hat{w} \times \hat{h}$, which 1) maintains good shape preservation for those spatially important content; 2) adapts new viewing configuration with adjustable depth; and 3) leaves the space for user to edit the important content for interactive applications. Based on these considerations, we formulate stereoscopic image resizing as an optimization

problem to find a set of deformed meshes $\hat{\mathbf{V}}$ and adjusted depths $\hat{\mathbf{Z}}$, which minimize the following objective functions:

$$E(\hat{\mathbf{V}}, \hat{\mathbf{Z}}) = E_Q(\hat{\mathbf{V}}) + \lambda_S E_S(\hat{\mathbf{V}}, \hat{\mathbf{Z}}) + \lambda_I E_I(\hat{\mathbf{V}}, \hat{\mathbf{Z}}) \quad (1)$$

where $E_Q(\hat{\mathbf{V}})$ is the image quality energy, $E_S(\hat{\mathbf{V}}, \hat{\mathbf{Z}})$ is the stereoscopic quality energy, $E_I(\hat{\mathbf{V}}, \hat{\mathbf{Z}})$ is the important content energy, and λ_D and λ_I are the weights for the corresponding energy terms. These energy terms well explain the above three requirements for stereoscopic image retargeting. In what follows, we will discuss the energy terms in detail.

B. IMAGE QUALITY ENERGY

Refer to the existing warping-based retargeting methods [8], [9], we also evaluate quality from two aspects: *shape distortion* and *line bending*. The first one measures the dissimilarity between the deformed and its original shapes to retain the geometry structure, and the second one measures the angle between the deformed and its original edges to minimize the bending of the grid edges. Therefore, the image quality energy E_Q is defined as

$$E_Q(\hat{\mathbf{V}}) = E_{SD}(\hat{\mathbf{V}}) + \lambda_{LB} E_{LB}(\hat{\mathbf{V}}) \quad (2)$$

where E_{SD} is the shape distortion energy, E_{LB} is the line bending energy, and λ_{LB} is the weight for E_{LB}.

Since we employ grid-based warping to preserve the visually important content with small deformation while allow large deformation for those unimportant content, refer to [9], we adopt a similarity transformation to evaluate the shape distortion energy in resizing. For each mesh with four vertices (e.g., $\mathbf{V}_k$), the shape distortion energy is defined as

$$E_{SD}(k) = \sum_{i=1}^{4} \left\| \rho_k(\mathbf{v}_k^i) - \hat{\mathbf{v}}_k^i \right\|^2 \quad (3)$$

where ρ_k represents the similarity transformation for each grid mesh. The similarity transformation (only containing scale and translation) can be formulated as:

$$\rho_k(\mathbf{v}_k^i) = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} x_k^i \\ y_k^i \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (4)$$

With (4), minimization of (3) is equivalent to solving a linear least-squares problem $\mathbf{AP} = \mathbf{b}$, where

$$\mathbf{A}_k = \begin{bmatrix} x_k^1 & 0 & 1 & 0 \\ 0 & y_k^1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_k^4 & 0 & 1 & 0 \\ 0 & y_k^4 & 0 & 0 \end{bmatrix} \quad (5)$$

$$\tilde{\mathbf{b}}_k = \begin{bmatrix} \hat{x}_k^1 & \hat{y}_k^1 & \cdots & \hat{x}_k^4 & \hat{y}_k^4 \end{bmatrix}^T \quad (6)$$

It is straightforward to obtain the similarity transform matrix with the form as follows:

$$\mathbf{P}_k = (\mathbf{A}_k^T \mathbf{A}_k)^{-1} \mathbf{A}_k^T \hat{\mathbf{b}}_k \quad (7)$$

It should be noted that the rotation factor in the similarity transformation is ignored, because only the horizontal camera array is considered in the retargeted image. Ideally, if a mesh undergoes a similarity transformation, the expected position $\tilde{\mathbf{b}}_k = \mathbf{A}_k \mathbf{P}_k$ should be identical to $\hat{\mathbf{b}}_k$ (i.e., the position of $\mathbf{v}_k$ after deformation). Thus, the shape distortion energy can be calculated by summing $\left\| \tilde{\mathbf{b}}_k - \hat{\mathbf{b}}_k \right\|^2$ for all meshes as follows:

$$E_{SD}(\hat{\mathbf{V}}) = \sum_{\mathbf{V}_k \in \mathbf{V}} S(k) \cdot \left\| \tilde{\mathbf{b}}_k - \hat{\mathbf{b}}_k \right\|^2 \quad (8)$$

where $S(k)$ is the significance value of mesh k, which is computed using the 3D saliency detection approach in our previous work [38]. The goal of the importance weighting is to preserve shapes in those high-significance objects (with high weight). To avoid undue deformation, the significance value is normalized to [0, 1]. Thus, by involving the significance values, minimizing E_{SD} will preserve the shape of important content while distribute more distortion to the unimportant content.

Refer to [7], to minimize the bending of the mesh edges, we define a line bending energy by measuring the angle between the deformed edge and its original edge. Let $\mathbf{e}_{i,j} = \mathbf{v}_i - \mathbf{v}_j$ and $\hat{\mathbf{e}}_{i,j} = \hat{\mathbf{v}}_i - \hat{\mathbf{v}}_j$ as edges, the angle between $\mathbf{e}_{i,j}$ and $\hat{\mathbf{e}}_{i,j}$ is approximated by

$$E_{LB}(\hat{\mathbf{e}}_{i,j}) = \left\| (\hat{\mathbf{v}}_i - \hat{\mathbf{v}}_j) - s_e(\mathbf{v}_i - \mathbf{v}_j) \right\|^2 \quad (9)$$

where s_e is a scale parameter. Refer to the solution in (7), minimization of (9) yields a linear least-squares solution as

$$\Delta(\hat{\mathbf{e}}_{i,j}) = \left\| \mathbf{e}_{i,j} (\hat{\mathbf{e}}_{i,j}^T \mathbf{e}_{i,j})^{-1} \mathbf{e}_{i,j}^T \hat{\mathbf{e}}_{i,j} - \hat{\mathbf{e}}_{i,j} \right\|^2 \quad (10)$$

Let $\mathbf{C}_{\mathbf{e}_{i,j}} = \mathbf{e}_{i,j} (\hat{\mathbf{e}}_{i,j}^T \mathbf{e}_{i,j})^{-1} \mathbf{e}_{i,j}^T - \mathbf{I}$, the line bending energy for all meshes is defined as

$$E_{LB}(\hat{\mathbf{V}}) = \sum_{<\mathbf{v}_i, \mathbf{v}_j> \in \mathbf{E}} \left\| \mathbf{C}_{\mathbf{e}_{i,j}} \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{v}}_i \\ \hat{\mathbf{v}}_j \end{bmatrix} \right\|^2 \quad (11)$$

C. DEPTH QUALITY ENERGY

As depth ranges of different display devices will be not the same, the goal of depth adaptation here is to adapt the shape and perspective of the retargeted content. According to perspective projection, the scale of a stereoscopic object depends on the depth. Thus, when there are large depth changes due to depth adaption, the scale of the object must be adjusted according to the adapted depths to match such changes. Instead of preserving the original depth/disparity information usually in the previous works [16]–[18], we seek to derive the relationship between the shape scaling and the perceived depth so that we can scale the object adaptively according to the target depth. The relationship between the perceived depth and the disparity is determined as follows:

$$Z_p = \frac{d_e}{d_e + (x_p^L - x_p^R)} \cdot L_D \quad (12)$$

where x_p^L and x_p^R are the projections of a point p on the left and right images, and $d_p = x_p^L - x_p^R$ denotes the disparity.

L_D is the viewing distance from the viewers to the screen, and d_e is the interocular distance between the viewer's two eyes. Since depth perception will be affected by different viewing distances and display devices if displaying the retargeted content on different displays, we focus on controlling the perceived depth instead of the disparity for better 3D viewing in this paper. Here, we use the stereo matching algorithm in [39] to estimate the disparity.

Since the convergence/accommodation limitations on depth range for different display devices will be not the same, to provide proper depth perception on different display types and resolutions, user favorite depth range (influenced by retinal disparity limits and viewing distance) is defined to adjust the depth range to avoid accommodation-convergence conflict. Considering the factors of viewing distance and display device, the target depth range for a particular viewing configuration can be determined as follows:

$$\hat{Z}_{\max} = \frac{d_e \hat{L}_D}{d_e - \eta_1 \hat{L}_D}, \hat{Z}_{\min} = \frac{d_e \hat{L}_D}{d_e - \eta_2 \hat{L}_D} \tag{13}$$

where η_1 and η_2 denote the negative and positive retinal disparity limits, respectively, and $\hat{L}_D$ denotes the new viewing distance. These factors are highly related to the viewing configuration for the new retargeted content. Then, the transformation can be formulated as:

$$\begin{aligned} \hat{Z}_{\mathbf{v}_i} &= \frac{\hat{Z}_{\max} - \hat{Z}_{\min}}{Z_{\max} - Z_{\min}} \cdot (Z_{\mathbf{v}_i} - Z_{\min}) + \hat{Z}_{\min} \\ &= s_z \cdot Z_{\mathbf{v}_i} + t_z \end{aligned} \tag{14}$$

Obviously, (s_z, t_z) denotes the scale and translation factors for depth, which is similar with the definition of similarity transform in Eq. (4). In fact, by changing the viewing distance, the disparity (depth) between the left and right images will be changed. To match such viewing configuration, the shape of an object should be scaled correspondingly, e.g., large size for small viewing distance. Given the relationship, we calculated the perceived depths $\{Z_{\mathbf{v}_i}\}$ and $\{\hat{Z}_{\mathbf{v}_i}\}$ for all the vertices of the meshes before and after depth adaption. The scaling factor for size changing is defined by

$$k_{\mathbf{v}_i} = \frac{Z_{\mathbf{v}_i}}{\hat{Z}_{\mathbf{v}_i}} \tag{15}$$

It is obvious that the retinal disparity limits and viewing distance determine the scaling factor. If we retarget a 3D content produced for larger screens and display it on a small screen, the value of $k_{\mathbf{v}_i}$ is typically larger than 1. In this case, the size of an object should be down-scaled to adapt the target depth. To adjust the depth in a more nature way, the size of a mesh should be changed along with the depth [40]. Thus, the edge of the meshes should be adjusted synchronously with the depth of the meshes. For each edge in the mesh, the expected scaling factor $k_{\mathbf{e}_{ij}}$ is given by $(k_{\mathbf{v}_i} + k_{\mathbf{v}_j})/2$. Hence, we have the depth-dependent scaling factor to control the overall depth quality, which is composed of two shape scaling and depth control terms:

$$\begin{aligned} E_S(\hat{\mathbf{V}}, \hat{\mathbf{Z}}) &= E_{SC}(\hat{\mathbf{V}}) + E_{DC}(\hat{\mathbf{Z}}) \\ &= \sum_{(\mathbf{v}_i, \mathbf{v}_j) \in \mathbf{E}} \left\| (\hat{\mathbf{v}}_i - \hat{\mathbf{v}}_j) - k_{\mathbf{e}_{ij}} (\mathbf{v}_i - \mathbf{v}_j) \right\|^2 \\ &\quad + \lambda_{DC} \cdot \sum_{\mathbf{v}_i \in \mathbf{V}} \omega(Z_{\mathbf{v}_i}) \cdot \left\| \hat{Z}_{\mathbf{v}_i} - (s_z \cdot Z_{\mathbf{v}_i} + t_z) \right\|^2 \end{aligned} \tag{16}$$

where λ_{DC} denotes the weight for depth control term. Here, instead of directly computing the depth distortion energy, the influence of viewing distance is integrated into the depth distortion energy, so that the scene near the screen will have large space to adjust its depth, which is defined as:

$$\omega(Z_{\mathbf{v}_i}) = \exp\left(\frac{|Z_{\mathbf{v}_i} - L_D|}{Z_{\max} - Z_{\min}} \right) \tag{17}$$

D. IMPORTANT CONTENT ENERGY

With only the image quality and depth quality energy terms, the optimal solution can be found to preserve the shape and depth constraints. However, in many interactive applications, the users expect to freely select the viewing content, and adjust the size and depth range interactively [41]. Therefore, we incorporate a new important content energy term, which provides more space for user to selectively present important content for interactive applications. The important content energy should satisfy the following two requirements: 1) other information of the image outside the selected important content should also be preserved for the completeness of content retargeting; 2) this term should be naturally integrated into the warping framework, i.e., it should be not isolated with the above global image quality and depth quality optimization.

To achieve the above goal, we draw a spatial cropping window as important content selected with the highest significance value. Of course, the users can manually select the important content based on users' own preference. For the meshes located within the window, they will not follow the same scaling factors in the above image quality and depth quality energy terms, but are assigned user specified scaling factors for this important content. Let $x_{i,j}$ denotes the horizontal coordinate of a mesh vertex in the i-th row and the j-th column, and $Z_{i,j}$ denotes the corresponding depth of the mesh vertex. Only considering the scaling on the horizontal and depth directions, we introduce two user specified scaling factors (s'_x and s'_z) to construct the energy term, which is composed of two object scaling and depth scaling terms:

$$\begin{aligned} E_I(\hat{\mathbf{V}}, \hat{\mathbf{Z}}) &= E_{OS}(\hat{\mathbf{V}}) + E_{DS}(\hat{\mathbf{Z}}) \\ &= \sum_{(\mathbf{v}_i, \mathbf{v}_j) \in \mathbf{E}^R} \left\| (\hat{x}_{i,j+1} - \hat{x}_{i,j}) - s'_x (x_{i,j+1} - x_{i,j}) \right\|^2 \\ &\quad + \lambda_{DS} \cdot \sum_{\mathbf{v}_i \in \mathbf{V}^R} \left\| \hat{Z}_{\mathbf{v}_i} - s'_z Z_{\mathbf{v}_i} \right\|^2 \end{aligned} \tag{18}$$

where λ_{DS} is the weight for depth scaling term, and $\mathbf{E}^R$ and $\mathbf{V}^R$ means the set of edges and meshes on the selected region, respectively. Particularly, if $s'_x = 1$, this term ensures the important content to have the same size with the original object, and to display with sufficient space. Larger s'_x and s'_z will yield stronger 3D viewing of the selected content. The essence of the optimization is to effectively distribute the distortions across different regions no matter inside or outside the important content. Therefore, E_C is also resolved by a global solution.

Since user's input and interaction is an important function in the framework, especially in the above important content energy term, the user can specify a window to define an object. For image retargeting, the user is prompted to provide the target scaling factors s'_x and s'_z for the specified object in the x and z directions respectively. By doing so, the user can choose what he is interested in and display it with different freedom within the retargeted image. With larger horizontal scaling factors, the non-important contents will be discarded or encouraged larger distortion due to the global optimization essence. For the image containing multiple visually important regions, selection of different clipping windows will provide different descriptions for the scene.

E. OPTIMIZATION

The three energy terms are necessary to achieve satisfying retargeting results. As analyzed above, $E_Q(\hat{\mathbf{V}})$ is used to control the shape of the global meshes, $E_S(\hat{\mathbf{V}}, \hat{\mathbf{Z}})$ is used to constrain the depth range and scale of the global meshes, and $E_I(\hat{\mathbf{V}}, \hat{\mathbf{Z}})$ is used to determine the size and depth range for the user-specified important content. Minimizing the total energy function corresponds to solving a least-squares linear system ($\mathbf{AP} = \mathbf{b}$), and a set of deformed vertices $\hat{\mathbf{V}}$ and retargeted depths $\hat{\mathbf{Z}}$ for all vertices can be found. Then, the adjusted depth is converted back to obtain the retargeted disparity as

$$\hat{d}_i = d_e \cdot \left(1 - \frac{L_D}{\hat{Z}_i}\right) \tag{19}$$

Based on the original and the retargeted disparities, $\mathbf{V}$ and $\hat{\mathbf{V}}$ are projected to the right image, obtaining the original meshes $\mathbf{V}_R$ and the deformed meshes $\hat{\mathbf{V}}_R$, respectively. Thus, by establishing the mapping relationships between $\mathbf{V}$ and $\hat{\mathbf{V}}$, and between $\mathbf{V}_R$ and $\hat{\mathbf{V}}_R$ respectively, we can map the left and right images to the target resolution while preserving the shapes and adjusting the depth range of important content in global and local semantics. Different with [16]–[18] that optimize the left and right images simultaneously, we use depth information from left image to construct the mesh set for the right image. However, if a mesh cannot find the correspondence in the other image (belonging to occluded/disoccluded regions), the adjacent meshes are extended to preserve the region's width in constraining the local warping function.

In the experiment, we find that even though we can resolve the solution only with $E_Q(\hat{\mathbf{V}})$, as shown in Fig.1(a), without $E_S(\hat{\mathbf{V}}, \hat{\mathbf{Z}})$, the depth range cannot be properly adjusted, which makes the performance of depth adaptation is limited, while without $E_I(\hat{\mathbf{V}}, \hat{\mathbf{Z}})$, the user interaction is not considered, thus the selected content is not changed in the retargeting result. We have tried many different weights for the energy combination in Eq. (1), and the results shows some differences. In general, we should place enough emphasis on $E_I(\hat{\mathbf{V}}, \hat{\mathbf{Z}})$ to produce prominent stereoscopic vision for the user-specified content, and too small value λ_I can make the term weak.

IV. EXPERIMENTAL RESULTS AND ANALYSES

In this section, we evaluate the performance of the propose method with other methods on the stereoscopic images with different depth ranges. We collect several stereoscopic images with large depth ranges from IVY Lab Stereo 3D Image Database for Disparity Remapping [42], and with relatively small depth ranges from PSU Stereo Saliency Benchmark [43]. To evaluate important content, all these testing images contain independent salient objects. The representative images are shown in Fig. 3.

FIGURE 3. The selected testing images in the experiment.

In the experiment, we set the interocular distance $d_e = 65$mm, the viewing distance $L_D = 1200$mm, and the disparity limits $\eta_1 = -1$ and $\eta_2 = 1$. The parameters L_D and (η_1, η_2) are adjustable and slight changes above or below the values will have certain effects on the retargeting performance. Also, there are other adjustable parameters, such as λ_{LB}, λ_{DC}, λ_{DS}, λ_S, λ_I, s'_x and s'_z. Through comprehensively comparison from accuracy, comfortability and controllability, we set $\lambda_{LB} = 1.25$, $\lambda_{DC} = 0.25$, $\lambda_{DS} = 0.025$, $\lambda_S = 1.5$ and $\lambda_I = 1.25$ to achieve satisfying results, and $s'_x = 1$, $s'_z = 1$ to simply preserve the original size and depth for the important content as much as possible. Since many parameters are adjustable in our method, our results may be not remarkable in a single indictor (e.g., accuracy, comfortability or controllability) for some testing images, but it provides a distinctive solution for content retargeting, depth adaptation and interactive editing simultaneously to maximize the user's experience.

A. QUALITATIVE COMPARISON WITH OTHER METHODS

We compare our method with three state-of-the-art stereoscopic image retargeting methods, including single-layer

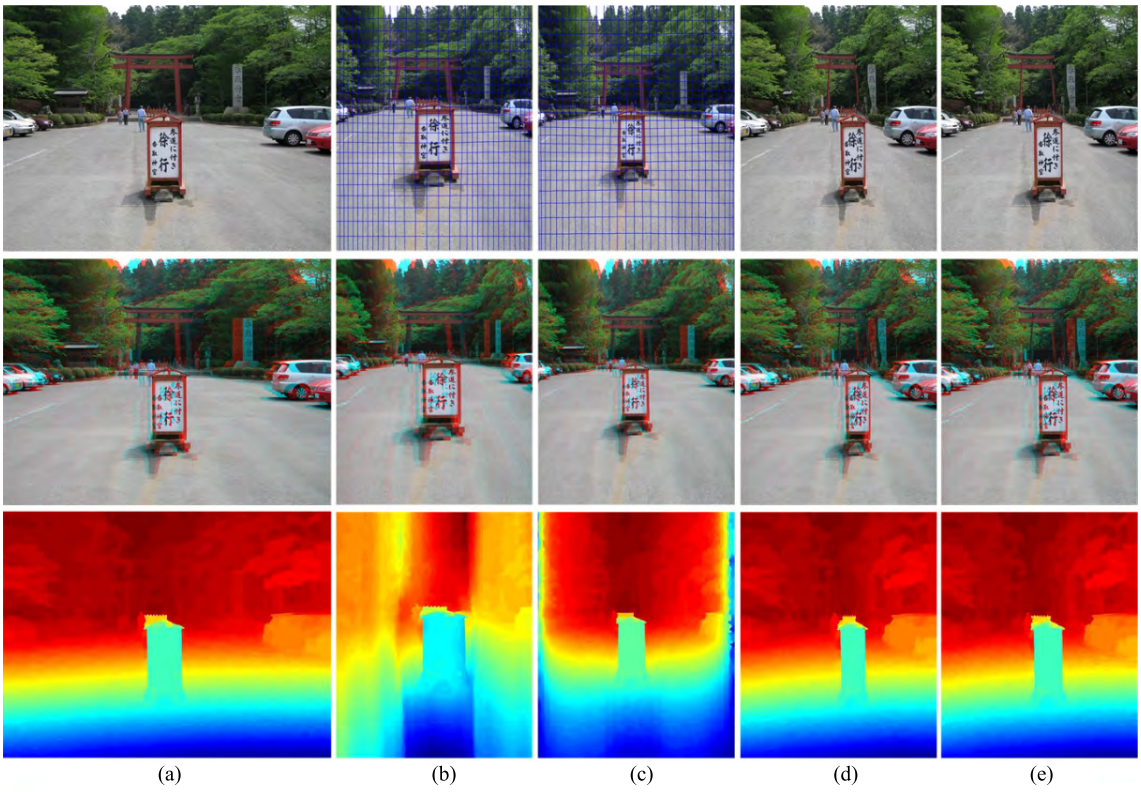


FIGURE 4. Results on testing image #7. From top to bottom: left images, the retargeted stereoscopic images (shown in red-cyan anaglyph), and the disparity maps (shown in pseudo-color map). (a) Original image. (b) Ours. (c) SLWAR. (d) GCSSC. (e) VASSC.

warping-based retargeting approach (SLWAR) [16], geometrically consistent stereoscopic seam carving approach (GCSSC) [14], and visual attention guided stereoscopic seam carving (VASSC) approach [20] (our previous work). A qualitative comparison of the retargeting results on two stereoscopic images with relatively small depth range are presented in Figs. 4-5. All images are shrunk by 40%. From the figure, we find that the seam carving based retargeting methods (GCSSC and VASSC) will lead to subtle shape deformation, e.g., the edge of the foreground signboard and the background archway in Figs. 4 (d) and (e), and the legs of the sculpture in Figs. 5 (d) and (e). For the SLWAR method, due to disparity preservation and warping optimization essence, the objects are retargeted to a small size with original depth range. In contrast, since the depth quality energy and important energy are simultaneously considered, our method not only preserves the original size of the important object (from the viewpoint of the retargeted image, the size of the object is expanded), but also enhances the depth sensation for those test image pairs with small depth range (the depth range is adaptively expanded). The phenomenon is very obvious observed from the disparity maps in Figs. 4-5.

In addition, we also illustrate the retargeting results on two stereoscopic images with large depth range in Figs. 6-7 (the stereoscopic images are visually uncomfortable). Similar with the above conclusion, the GCSSC and VASSC methods will deform the shapes of cup in Figs. 6 (d) and (e) and kettle in Figs. 7 (d) and (e). More importantly, the retargeting results still have large depth range outside the CVZ, leading to poor visual comfort experience. The SLWAR method will result in small object size and large depth range for the testing images in Figs. 6-7. In conclusion, the seam carving methods may lead to shape deformation due to irregular seams, and the warping-based method may lead to object size reduction. All these methods enforce directly or indirectly disparity/depth preservation constraint. The proposed method takes the depth quality energy and important energy into account, and achieves better user's visual experience.

B. THE INFLUENCE OF VIEWING CONFIGURATION

From the perspective of depth adaptation, if a retargeted stereoscopic image is displayed on a small screen, its viewing distance should be reduced correspondingly. In our depth quality energy, we control the depth in the 3D space to allow user to edit the viewing distance. In Fig. 8, we set three viewing distances ($D = 800$, 1000 and 1200) for four scaling ratios (50%, 40%, 30%, and 20%), respectively. We can see that, by adding the viewing distance, the depth difference between objects is slightly changed, i.e., the relative depth distance between objects is increased. This phenomenon is understandable because with the increased viewing distance,

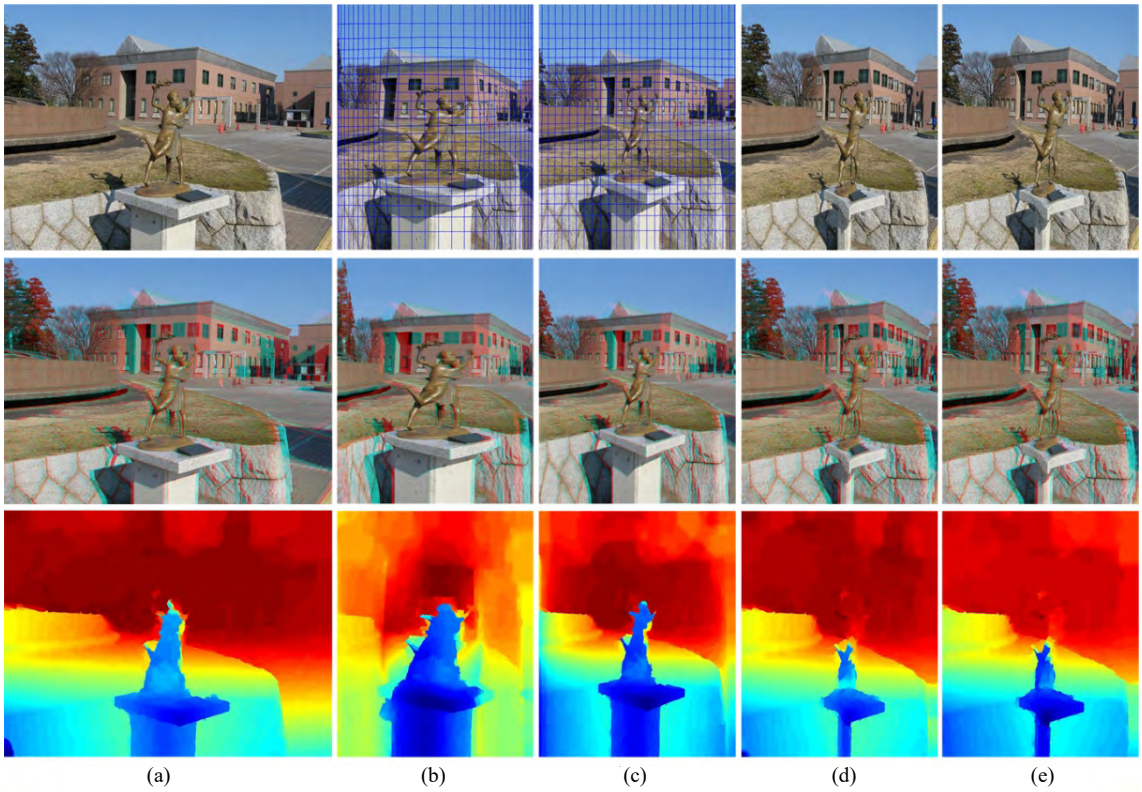


FIGURE 5. Results on testing image #8. From top to bottom: left images, the retargeted stereoscopic images (shown in red-cyan anaglyph), and the disparity maps (shown in pseudo-color map). (a) Original image. (b) Ours. (c) SLWAR. (d) GCSSC. (e) VASSC.

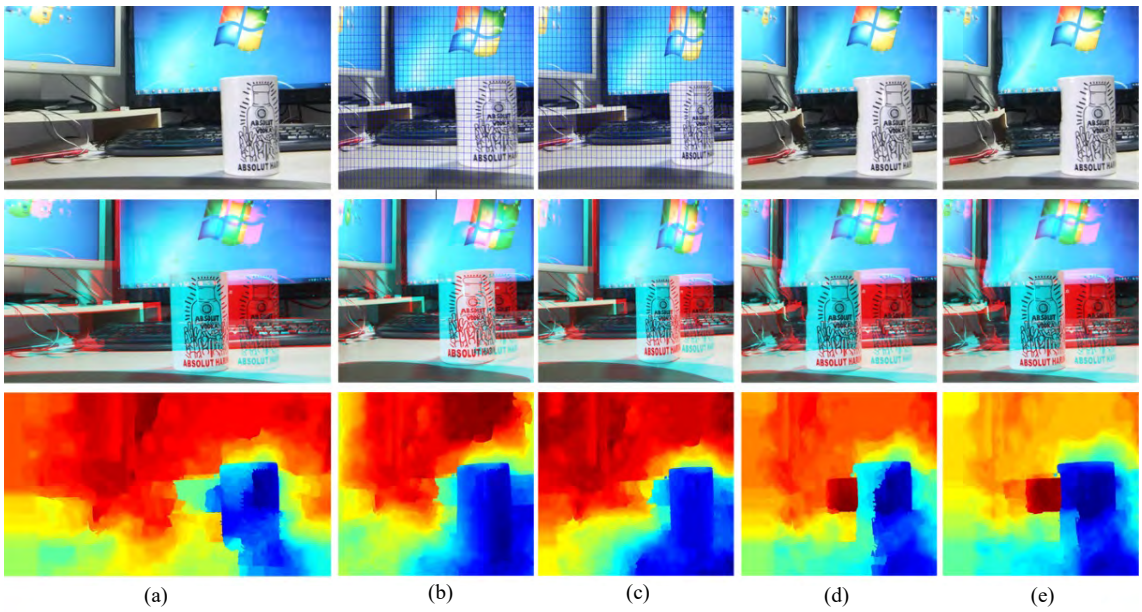


FIGURE 6. Results on testing image #3. From top to bottom: left images, the retargeted stereoscopic images (shown in red-cyan anaglyph), and the disparity maps (shown in pseudo-color map). (a) Original image. (b) Ours. (c) SLWAR. (d) GCSSC. (e) VASSC.

the viewing perspective is also increased accordingly. In the experiment, since the range of the selected viewing distances is not very large, their influences on depth perception in the result are not obvious. Normally, the viewing distance is set to three times the height of display. Thus, if we display the retargeted stereoscopic images on different displays, the

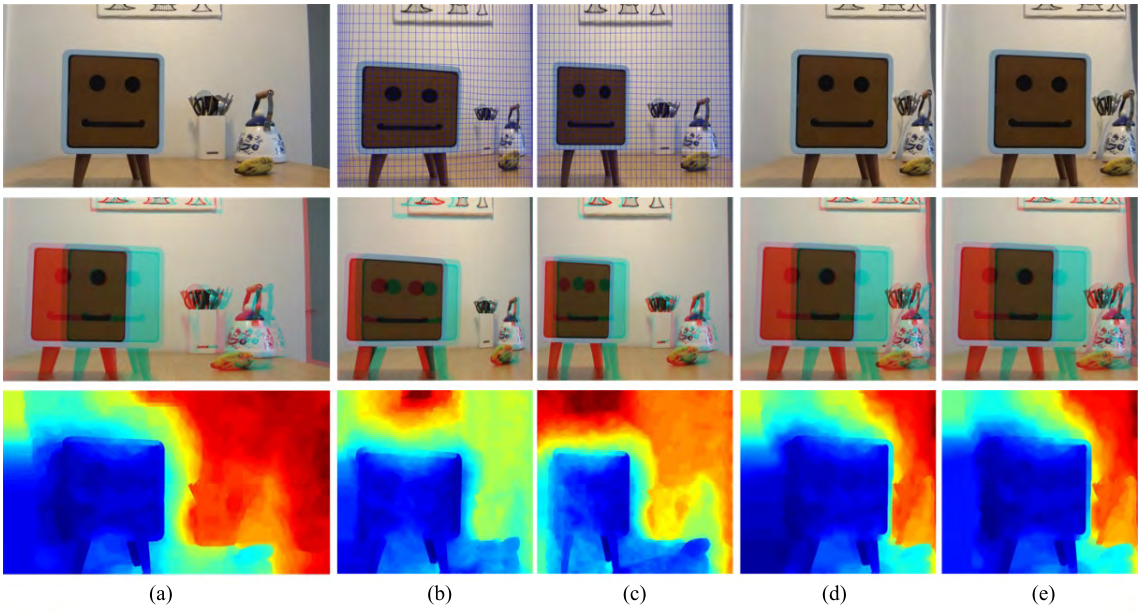


FIGURE 7. Results on testing image #2. From top to bottom: left images, the retargeted stereoscopic images (shown in red-cyan anaglyph), and the disparity maps (shown in pseudo-color map). (a) Original image. (b) Ours. (c) SLWAR. (d) GCSSC. (e) VASSC.

FIGURE 8. Results of different viewing distances for shrinking the image width by: 50%, 40%, 30% and 20%.

viewing distance should be carefully adjusted to provide a better depth adaptation.

C. THE INFLUENCE OF DIFFERENT ENERGY TERMS

In addition, to better demonstrate the impact of each energy term in the proposed retargeting framework, we design the following four schemes for comparison. There are three major energy terms E_Q, E_S and E_I in our formulation, including six sub-energy terms. We design the following schemes listed in Table 1, denoted by Scheme-1, Scheme-2, Scheme-3 and Proposed scheme. The same E_Q is included in these schemes, and other optional component (e.g., shape scaling,

FIGURE 9. The retargeting results with different energy combination schemes. (a) Original image. (b) Scheme-1. (c) Scheme-2. (d) Scheme-3. (e) Proposed.

depth control, object scaling or depth scaling) is added to construct these schemes. The retargeting results are shown in Fig. 8. Three important conclusions can be drawn from the results: 1) without $E_{OS}(\hat{\mathbf{V}})$, the objects are retargeted to a small size; 2) compared with Scheme-2 and Scheme-3 with or without $E_{DS}(\hat{\mathbf{Z}})$, the local depth scaling will affect the global depth layout; 3) compared with Scheme-3 and Proposed scheme with or without $E_{SC}(\hat{\mathbf{V}})$, by integrating the influence of the perceived depth, the size of an object is adaptively scaled, to avoid the important content missing as shown clearly in the last row of Fig. 8. Therefore, we can conclude that cooperation of these energy terms will provide more natural subjective perception for users.

D. USER STUDY

We also perform user study to assess our algorithms. 20 participants were participated in our user studies. We conduct subjective experiment on a Samsung UA65F9000 65 inch Ultra HD 3D-LED TV with 3D shutter glasses. We perform paired comparisons between the retargeted results obtained by our method and one of the comparative methods: SLWAR [16], GCSSC [14] and VASSC [20]. In the test,

TABLE 1. List of the schemes compared in this study.

Scheme	Description
Scheme-1	$\left\{E_Q(\hat{\mathbf{V}}), E_{DC}(\hat{\mathbf{Z}})\right\}$
Scheme-2	$\left\{E_Q(\hat{\mathbf{V}}), E_{DC}(\hat{\mathbf{Z}}), E_{OS}(\hat{\mathbf{V}})\right\}$
Scheme-3	$\left\{E_Q(\hat{\mathbf{V}}), E_{DC}(\hat{\mathbf{Z}}), E_{OS}(\hat{\mathbf{V}}), E_{DS}(\hat{\mathbf{Z}})\right\}$
Proposed	$\left\{E_Q(\hat{\mathbf{V}}), E_S(\hat{\mathbf{V}}, \hat{\mathbf{Z}}), E_I(\hat{\mathbf{V}}, \hat{\mathbf{Z}})\right\}$

9 stereoscopic image pairs in Fig. 3 are randomly chosen for this test. Each participant is asked to choose one retargeted result preferred to other according to the overall visual experience in terms of accuracy, comfortability and controllability, and performs 27 comparisons in each iteration. The user study results are reported in Table 2. These comparison results confirm the strong preference of our method over other methods. Specifically, our method receives 64.4% votes when compared to SLWAR, 73.9% votes when compared to GCSSC, and 79.4% votes when compared to VASSA. The preference is much more remarkable for those #1, #2, #3 and #4 testing

TABLE 2. The User's preference results of subjective paired comparisons.

Method	#1	#2	#3	#4	#5	#6	#7	#8	#9	Avg
Our vs SLWAR	65%	75%	60%	60%	55%	65%	90%	55%	55%	64.4%
Our vs GCSSC	95%	80%	75%	70%	75%	65%	80%	75%	50%	73.9%
Our vs VASSC	100%	75%	100%	90%	90%	50%	75%	70%	65%	79.4%

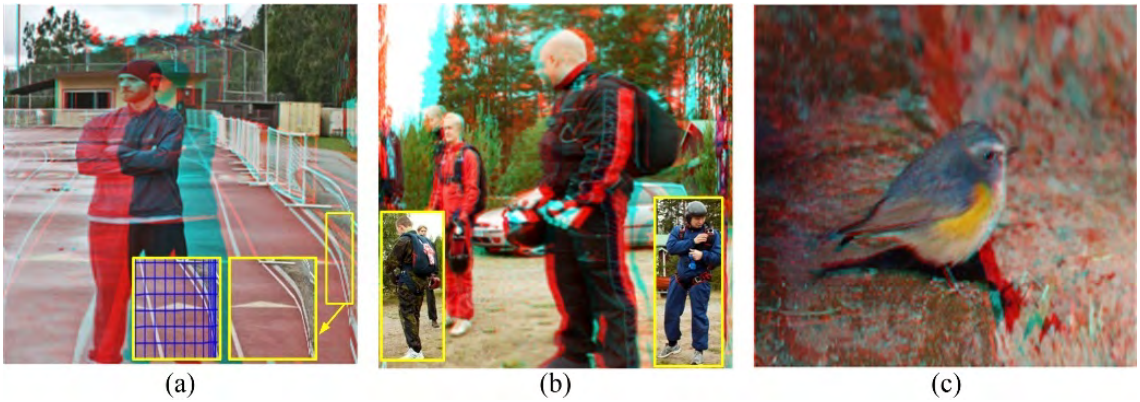


FIGURE 10. Failure cases of the proposed method.

images with large depth range. In general, participants favor our results over those of existing stereoscopic retargeting methods (e.g., SLWAR, GCSSC, and VASSC) due to ignoring user's experiences.

E. LIMITATIONS

The proposed approach has several limitations. First, when the captured scene contains slanted texture at the leftmost or rightmost side, example shown in Fig. 10 (a), our method does not work well. Due to the depth discontinuity between the corresponding and occluded regions (still using the original depth values for the occluded regions), there is significant ghost in the connection area, even though we extend the adjacent meshes to cover the occluded regions.

In addition, for the scene included multiple objects, if the selected important content is adjusted with a large scaling factors s'_x, it can lead to visual content missing in the retargeted image. Fig. 10 (b) shows such an example, where other two peoples on the left and right sides are discarded for the case of small scaling ratio. However, our method often works well for simple scene that do not contain multiple salient objects or strong textures at the background (e.g. in Fig. 10(c)).

Finally, our method relied on the weights to balance the distortions among the salient and important regions in a global image, and also to balance the local and global depth distributions. The same weight setting in fact cannot well for all interactive selections from drawing window, scaling factors and viewing distance. Therefore, it is needed to specify such weight setting, and more advanced weighting can be potentially employed for creating more accurate results.

V. CONCLUSIONS

In this paper, we present a content retargeting and depth adaptation method for stereoscopic image. By exploiting the complementarity relationship among image quality, depth quality and important content, we propose a unified grid optimization framework to fuse the three energy constraints. Besides, various user controllable retargeting scenarios are integrated to guide the optimization procedure. As a result, our system yields a retargeted result with high visual experience for users and shows a strong flexibility to process users' requirement, achieving a better trade-off among accuracy, comfortability and controllability for information presentation. As future work, we plan to consider additional aesthetics, object features, and more intelligent display-adaptive content retargeting [44]. We also plan to extend this framework to enable new applications, such as stereoscopic image stitching [45] and disparity manipulation [46].

REFERENCES

[1] O. Wang, M. Lang, M. Frei, A. Hornung, A. Smolic, and M. Gross, "StereoBrush: Interactive 2D to 3D conversion using discontinuous warps," in *Proc. 8th Eurographics Symp. Sketch-Based Interfaces Modeling*, 2011, pp. 47–54.

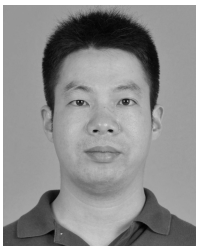
[2] S.-J. Luo, Y.-T. Sun, I.-C. Shen, B.-Y. Chen, and Y.-Y. Chuang, "Geometrically consistent stereoscopic image editing using patch-based synthesis," *IEEE Trans. Vis. Comput. Graphics*, vol. 21, no. 1, pp. 56–67, Jan. 2015.

[3] 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.

[4] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," *ACM Trans. Graph.*, vol. 26, no. 3, 2007, Art. no. 118,

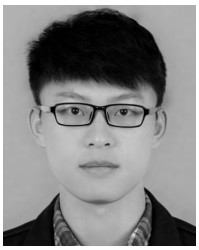
[5] M. Rubinstein, A. Shamir, and S. Avidan, "Improved seam carving for video retargeting," *ACM Trans. Graph.*, vol. 27, no. 16, pp. 1–9, 2008.

[6] M. Grundmann, V. Kwatra, M. Han, and I. Essa, "Discontinuous seam-carving for video retargeting," in *Proc. IEEE CVPR*, Jun. 2010, pp. 569–576.
[7] 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.
[8] Y. Guo, F. Liu, J. Shi, Z.-H. Zhou, and M. Gleicher, "Image retargeting using mesh parametrization," *IEEE Trans. Multimedia*, vol. 11, no. 5, pp. 856–867, Aug. 2009.
[9] G.-X. Zhang, M.-M. Cheng, S.-M. Hu, and R. R. Martin, "A shape-preserving approach to image resizing," *Comput. Graph. Forum*, vol. 28, no. 7, pp. 1897–1906, 2009.
[10] Y. Niu, F. Liu, W.-C. Feng, and H. Jin, "Aesthetics-based stereoscopic photo cropping for heterogeneous displays," *IEEE Trans. Multimedia*, vol. 14, no. 3, pp. 783–796, Jun. 2012.
[11] W. Wang, J. Shen, Y. Yu, and K.-L. Ma, "Stereoscopic thumbnail creation via efficient stereo saliency detection," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 8, pp. 2014–2027, Aug. 2017.
[12] S.-S. Lin, C.-H. Lin, Y.-H. Kuo, and T.-Y. Lee, "Consistent volumetric warping using floating boundaries for stereoscopic video retargeting," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 5, pp. 801–813, Mar. 2016.
[13] K. Utsugi, T. Shibahara, T. Koike, K. Takahashi, and T. Naemura, "Seam carving for stereo images," in *Proc. 3DTV-Conf., True Vis.-Capture, Transmiss. Display 3D Video (3DTV-CON)*, Jun. 2010, pp. 1–4.
[14] T. D. Basha, Y. Moses, and S. Avidan, "Stereo seam carving a geometrically consistent approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 10, pp. 2513–2525, Oct. 2013.
[15] J. Lei, M. Wu, C. Zhang, F. Wu, N. Ling, and C. Hou, "Depth-preserving stereo image retargeting based on pixel fusion," *IEEE Trans. Multimedia*, vol. 19, no. 7, pp. 1442–1453, Jul. 2017.
[16] 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.
[17] K.-Y. Lee, C.-D. Chung, and Y.-Y. Chuang, "Scene warping: Layer-based stereoscopic image resizing," in *Proc. IEEE Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 49–56.
[18] B. Li, L.-Y. Duan, C.-W. Lin, T. Huang, and W. Gao, "Depth-preserving warping for stereo image retargeting," *IEEE Trans. Image Process.*, vol. 24, no. 9, pp. 2811–2826, Sep. 2015.
[19] 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.
[20] F. Shao, W. Lin, W. Lin, G. Jiang, M. Yu, and R. Fu, "Stereoscopic visual attention guided seam carving for stereoscopic image retargeting," *J. Display Technol.*, vol. 12, no. 1, pp. 22–30, Jan. 2016.
[21] Y. Fang, J. Wang, Y. Yuan, J. Lei, W. Lin, and P. Le Callet, "Saliency-based stereoscopic image retargeting," *Inf. Sci.*, vol. 372, pp. 347–358, Dec. 2016.
[22] Y. Chen, Y. Pan, M. Song, and M. Wang, "Improved seam carving combining with 3D saliency for image retargeting," *Neurocomputing*, vol. 151, pp. 645–653, Mar. 2015.
[23] S.-S. Lin, C.-H. Lin, S.-H. Chang, and T.-Y. Lee, "Object-coherence warping for stereoscopic image retargeting," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 5, pp. 759–768, May 2014.
[24] Y. Niu, W.-C. Feng, and F. Liu, "Enabling warping on stereoscopic images," *ACM Trans. Graph.*, vol. 31, no. 6, 2012, Art. no. 183.
[25] 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.
[26] M. B. Islam, L. K. Wong, C.-O. Wong, and K.-L. Low, "Stereoscopic image warping for enhancing composition aesthetics," in *Proc. ACPR*, Nov. 2015, pp. 645–649.
[27] Y. Liu, L. Sun, and S. Yang, "A retargeting method for stereoscopic 3D video," *Comput. Vis. Media*, vol. 1, no. 2, pp. 119–127, Jun. 2015.
[28] S. Kopf, B. Guthier, C. Hipp, J. Kiess, and W. Effelsberg, "Warping-based video retargeting for stereoscopic video," in *Proc. ICIP*, Oct. 2014, pp. 2898–2902.
[29] 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.
[30] 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.
[31] 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.
[32] M. Wang, X.-J. Zhang, J.-B. Liang, S.-H. Zhang, and R. R. Martin, "Comfort-driven disparity adjustment for stereoscopic video," *Comput. Vis. Media*, vol. 2, no. 1, pp. 3–17, Mar. 2016.
[33] W. Yan, C. Hou, B. Wang, and L. Wang, "Content-aware disparity adjustment for different stereo displays," *Multimedia Tools Appl.*, vol. 76, pp. 10465–10479, Apr. 2017. doi: 10.1007/s11042-016-3442-y.2016.
[34] H. Park, H. Lee, and S. Sull, "Efficient viewer-centric depth adjustment based on virtual fronto-parallel planar projection in stereo 3D images," *IEEE Trans. Multimedia*, vol. 16, no. 2, pp. 326–336, Feb. 2014.
[35] C. Oh, B. Ham, S. Choi, and K. Sohn, "Visual fatigue relaxation for stereoscopic video via nonlinear disparity remapping," *IEEE Trans. Broadcast.*, vol. 61, no. 2, pp. 142–153, Jun. 2015.
[36] F. Shao, W. Lin, Z. Li, G. Jiang, and Q. Dai, "Toward simultaneous visual comfort and depth sensation optimization for stereoscopic 3-D experience," *IEEE Trans. Cybern.*, vol. 47, no. 12, pp. 4521–4533, Dec. 2017.
[37] J. Wang *et al.*, "Adaptive content condensation based on grid optimization for thumbnail image generation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 11, pp. 2079–2092, Nov. 2016.
[38] Q. Jiang, F. Shao, G. Jiang, M. Yu, Z. Peng, and C. Yu, "A depth perception and visual comfort guided computational model for stereoscopic 3D visual saliency," *Signal Process., Image Commun.*, vol. 38, pp. 57–69, Oct. 2015.
[39] D. Sun, S. Roth, and M. J. Black, "Secrets of optical flow estimation and their principles," in *Proc. IEEE Intl. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2010, pp. 2432–2439.
[40] S.-J. Luo, I. 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, Nov. 2012, Art. no. 182.
[41] C.-H. Tan, M. B. Islam, L.-K. Wong, and K.-L. Low, "Semantics-preserving warping for stereoscopic image retargeting," in *Proc. Pacific-Rim Symp. Image Video Technol.*, 2005, pp. 257–268.
[42] H. Sohn, Y. J. Jung, S.-I. Lee, and Y. M. Ro. (2013) *IVY Lab Stereo 3D Image Database for Disparity Remapping*. [Online]. Available: http://ivylab.kaist.ac.kr/demo/3D/DisparityRemapping.htm
[43] Y. Niu, Y. Geng, X. Li, and F. Liu, "Leveraging stereopsis for saliency analysis," in *Proc. of IEEE CVPR*, Jun. 2012, pp. 454–461.
[44] B. Masia, G. Wetzstein, C. Aliaga, R. Raskar, and D. Gutierrez, "Display adaptive 3D content remapping," *Comput. Graph.*, vol. 37, no. 8, pp. 983–996, Dec. 2013.
[45] 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.
[46] P. Kellnhofer, P. Didyk, K. Myszkowski, M. M. Hefeeda, H.-P. Seidel, and W. Matusik, "GazeStereo3D: Seamless disparity manipulations," *ACM Trans. Graph.*, vol. 35, no. 4, Jul. 2016, Art. no. 68.

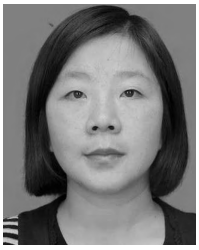


FENG SHAO (M'16) received the B.S. and Ph.D. degrees in electronic science and technology from Zhejiang University, Hangzhou, China, in 2002 and 2007, respectively. He was a Visiting Fellow with the School of Computer Engineering, Nanyang Technological University, Singapore, from 2012 to 2012. He is currently a Professor with the Faculty of Information Science and Engineering, Ningbo University, China. He has published over 100 technical articles in refereed journals and proceedings in the areas of 3D video coding, 3D quality assessment, and image perception. He received the Excellent Young Scholar Award by the National Natural Science Foundation of China, in 2016.

LIBO SHEN received the B.S. degree from the Nanjing University of Posts and Telecommunications, China, in 2016. He is currently pursuing the M.S. degree with Ningbo University, Ningbo, China. His current research interests include image/video processing and quality assessment.



QIUPING JIANG (S'17) received the Ph.D. degree from Ningbo University, Ningbo, China, in 2018, where he is currently an Associate Professor with the School of Information Science and Engineering. From 2017 to 2018, he was a Visiting Student with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research interests include image processing, visual perception modeling, and computer vision. He was a recipient of the JVCI 2017 Best Paper Award Honorable Mention as the first author. He is a Reviewer for several prestigious journals and conferences, such as the IEEE TNNLS, the IEEE TIP, the IEEE TCSVT, the IEEE TMM, the IEEE TSIPN, ICME, and ICIP.



FUCUI LI received the M.S. degree from the Hefei University of Technology, Hefei, China, in 2004. She is currently a Lecturer with the Faculty of Information Science and Engineering, Ningbo University, China. Her research interests mainly include 3D video and multimedia representation.

YO-SUNG HO (SM'06–F'16) received the B.S. and M.S. degrees in electronic engineering from Seoul National University, Seoul, South Korea, in 1981 and 1983, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California at Santa Barbara, in 1990. He joined the Electronics and Telecommunications Research Institute (ETRI), Daejon, South Korea, in 1983. From 1990 to 1993, he was with Philips Laboratories, Briarcliff Manor, NY, USA, where he was involved in the development of the advanced digital high-definition television system. In 1993, he rejoined the Technical Staff of ETRI and was involved in the development of the Korean DBS digital television and high-definition television systems. Since 1995, he has been with the Gwangju Institute of Science and Technology, Gwangju, South Korea, where he is currently a Professor of the Information and Communications Department. His research interests include digital image and video coding, image analysis and image restoration, advanced video coding techniques, digital video and audio broadcasting, three-dimensional video processing, and content-based signal representation and processing.