

A retargeting method for stereoscopic 3D video

Yi Liu¹ (✉), Lifeng Sun², and Shiqiang Yang²

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract We propose a disparity-constrained retargeting method for stereoscopic 3D video, which simultaneously resizes a binocular video to a new aspect ratio and remaps the depth to the perceptual comfort zone. First, we model distortion energies to prevent important video contents from deforming. Then, to maintain depth mapping stability, we model disparity variation energies to constraint the disparity range both in spatial and temporal domains. The last component of our method is a non-uniform, pixel-wise warp to the target resolution based on these energy models. Using this method, we can process the original stereoscopic video to generate new, high-perceptual-quality versions at different display resolutions. For evaluation, we conduct a user study; we also discuss the performance of our method.

Keywords stereoscopic 3D video; retargeting; disparity manipulation; image warping

1 Introduction

Stereoscopic 3D media has been available for a long time. With the development of mobile internet and social network technologies, stereoscopic videos are displayed on various devices with different display sizes. This brings the need for stereoscopic video retargeting techniques, to adapt videos to new aspect ratios.

In the traditional 2D broadcast world, fixed-window cropping and scaling are general methods

to fill top and bottom (or left and right) video areas. The same process is applied to each frame independently to achieve a new aspect ratio. However, due to the drawback of omitting some parts of the content, many content-aware methods have been proposed [1]. These preserve the region-of-interest from cropping and distortion. With the wide application of stereoscopic 3D techniques, research into content-aware retargeting methods has been carried out in the 3D image processing field. Existing methods for 2D media have been improved and applied to process 3D media.

However, differently from 2D media, the final visual experience is determined by both the image quality and stereo perception. Viewers see one image (frame) with the left eye and see the other with the right eye. Panum's area in the brain merges the two images to form a single image with relative depth perception [2]. The distance between a pair of corresponding points is called the disparity, and is an important depth cue [3]. Retargeting methods for 3D media must include methods for appropriate disparity processing.

Due to improper display configurations and image processing methods, when watching 3D videos, viewers often fail to see the stunning stereo visual experience that they deserve. Sometimes, viewers even feel visual fatigue or eye strain [4]. There are two main reasons for poor visual experience. The first one is the distortion of the image contents in 2D. Simple processing methods (e.g., scaling, cropping, uniform stretching) can not produce good video results in many cases because they lack some information from the original video [1]. The second reason is the arbitrary modification of the disparity in the stereoscopic field. A good stereo effect can be obtained by modifying the disparity properly according to the new view configuration.

1 Department of Computer Science and Technology, Tsinghua University, Tsinghua NLIST, Beijing 100084, China. E-mail: htcliuyi@163.com (✉).

2 Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China. E-mail: sunlf@tsinghua.edu.cn, yangshq@tsinghua.edu.cn.

Manuscript received: 2015-03-13; accepted: 2015-05-11

However, continual arbitrary disparity changes always cause viewers to feel fatigue [5]. Mismatched view configurations, e.g., screen distance, interocular distance, etc., can also cause discomfort. Thus, disparity values should be adjusted carefully in various scenes. Furthermore, compared to 3D still images, 3D video is more complicated because it involves time. Disparity constraints should be introduced not only in the spatial domain but also in the temporal domain. Thus the issue of stereoscopic 3D video retargeting is a more challenging problem. Key parts of the method must not only avoid content distortion, but also remap disparity to a proper range for each frame.

In this paper, we present a technique that extends existing warping-based 2D media retargeting methods to stereoscopic video. The objective of our work is to perform content-aware retargeting to the content of video frames and meanwhile adjust the corresponding disparity values for the stereoscopic video content. We first build energy models for the importance of the contents of the left and right frames. We try to preserve the more important parts in the left and right pictures with minimal distortion. Then we consider disparity constraints based on coherence in spatial and temporal domains, building appropriate energy models. Unlike previous methods, the proposed method considers both disparity values and disparity variation. It also considers objects' disparity relationships in video frames. Overall, we transform the 3D retargeting problem to a tradeoff of the various negative factors affecting 3D visual experience. In different application scenarios, we may alter weights for the energy components and minimize the sum of the distortion energies to obtain optimal results.

The main contributions of this paper can be summarized as follows:

- (1) We propose a new retargeting framework for stereoscopic 3D videos. The retargeting results can be adjusted according to various viewing scenarios.
- (2) We maintain disparity coherence in the temporal domain by constraining both the inter-frame disparity changes and the intra-frame disparity relation variation.
- (3) In the spatial domain, we use a linear depth remapping algorithm to manipulate the disparity range to lie in a comfort zone.

2 Related work

Previous literature has proposed many methods for retargeting a single 2D image while preserving its important content. Early methods intelligently crop away the surrounding content [6]. The seam carving method resizes images by removing or adding pixels judiciously [7, 8]. The warping method maps the original image to the target resolution according to some constraints [9–11], and is more robust for complicated pictures. In this paper, we improve the shape preservation model in Ref. [11].

Many of these single image resizing methods have been extended to 3D image (frame) resizing. Both Refs. [12] and [13] are based on the seam carving method, which sometimes impairs the image contents or changes their shape. Resizing systems based on the warping method are proposed in Refs. [14–16], while Refs. [17, 18] edit stereoscopic pairs by using a warping method. However, content representing structures such as lines is not preserved well. In Ref. [19], a cropping and warping method is used to improve the stereoscopic 3D experience. In Ref. [20], a layer-based resizing method is proposed. But it seems difficult to extend this method to stereo videos because of the uncertainty and inconvenience of layer segmentation.

For stereo videos, Ref. [21] focuses on disparity manipulation and proposes four disparity mapping operators to adjust disparity. However, depth relations of the adjacent features are not constrained. Besides, it does not support video resizing while adjusting depth, so some important problems in the resolution adaptation field, e.g., consistency of left and right pictures, are not considered by this method. A depth mapping method for stereo videos is proposed in Ref. [22]. To minimize distortion of stereoscopic image content, many energy models have been proposed to control the depth mapping process. However, the mesh edge preservation model is not a similar transformation constraint, and it can not ensure that shapes of image contents are maintained. Furthermore, coherence of rates of change of disparity is not constrained at key points, which can cause depth jitter. In Ref. [23], although the proposed system is fully automatic, the videos are cropped under 3D viewing constraints, risking loss of information. Like Ref. [3], the authors

of Ref. [24] state some principles in the field of stereoscopic media processing, which give us some inspiration when designing our disparity constraint models. Reference [25] conducts a series of eye-tracking experiments and tries to minimize the adaptation time for sudden temporal depth changes. Reference [26] focuses on the visual discomfort caused by disparity changes and uses a depth perception model to determine the time taken by the human visual system (HVS) to adapt to the changes.

3 The proposed method

We first detect the important parts in the key frames. In our experiments we use the saliency algorithm proposed in Ref. [27]. In order to incorporate high-level information, we also allow the viewer to optionally specify a region of interest. We then determine feature correspondences between the left and right frames for the disparity constraint. We extract SURF (speeded up robust features) using the method presented in Ref. [28] and perform a matching process between the two sets of key points. Estimating the fundamental matrix using the RANSAC method [29], we can get high quality matched key point pairs. Then, in order to control the disparity change in the temporal domain, we track the key point pairs across video key frames using the algorithm proposed in Ref. [30]. We construct quad meshes in left and right frames. To control the degree of impairment of the image quality and the stereo vision experience, we build energy models for various distortions. We minimize the total energy to preserve the significant information and control the disparity in the spatial and temporal domains. By computing the coordinates of the quads vertices, we transform every quad to generate a new picture with the target resolution and comfortable depth perception. An overview of our algorithm is given in Fig. 1. To demonstrate the function of our algorithm, we increase the horizontal disparity

while reducing the width of the image in the last two subfigures on the right. Details of the algorithm are described below.

3.1 Mesh deformation energy

Before modeling the energy, we compute the significance map for each picture. We consider three low-level visual features and integrate them to compute the significance at each pixel:

$$S = F_S \times F_G \times F_L \quad (1)$$

where F_S is the saliency map generated by the algorithm in Ref. [27], F_G is the gradient map of the image, F_L is the line segment map generated by the algorithm in Ref. [31], and S is the final significance map as shown in Fig. 1. If viewers specify the ROI (region of interest), the significant areas also can be propagated automatically into neighboring frames [32].

3.1.1 Grid deformation energy

As in Ref. [11], we attempt to ensure that important quads undergo a similarity transformation and compute the energy for every grid. We denote the i -th grid's energy as

$$E_g(i) = \sum_{k=1}^4 \|e'_k - s_f e_k\|^2 \quad (2)$$

where e'_k is the deformed version of the original k -th edge vector e_k . We compute the scale factor s_f using

$$\partial E_g(i) / \partial s_f = 0 \quad (3)$$

Unlike Ref. [11], we express the scale factor in terms of the eight unknown coordinate parameters of each quad's four vertices and substitute it into Eq. (2). Thus, instead of computing the scale factor value of every quad iteratively to approach the optimal solution, we obtain the optimal result quickly.

We define the grid deformation energy for both frames as

$$E_g = \sum_{i=1}^N E_{gL}(i) S_L(i) + \sum_{i=1}^N E_{gR}(i) S_R(i) \quad (4)$$

where N is the number of quads, E_{gL} is the energy

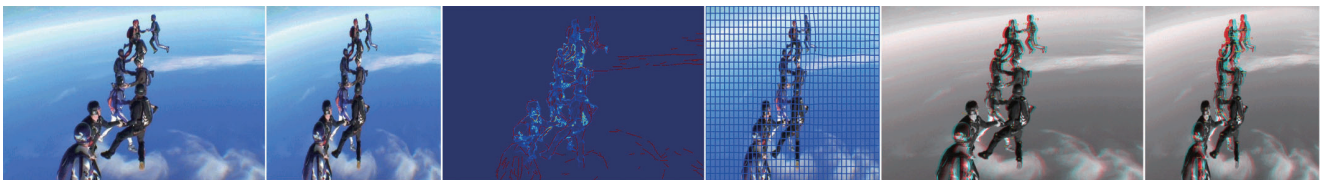


Fig. 1 Algorithm overview. Left to right: original left frame, output left frame, significance map of the left frame, warped mesh of the left frame, disparity map between the original left and right frame (o are key points in the left frame and + are those in the right frame), disparity map between the output left and right frame.

of the i -th grid in the left view, E_{gR} is the energy of the i -th grid in the right view, and $S_L(i)$ and $S_R(i)$ are the average significance values of all pixels in the i -th quad in the left and right frames respectively.

3.1.2 Edge distortion energy

Preservation of the important content, as well as the disparity constraints mentioned later, will sometimes seriously deform the grids in unimportant regions. We stop this happening by reducing the bending of the mesh edges. We define the bending energy for every edge in a frame as

$$E_e(i) = \|e'_i - s_e e_i\|^2 \tag{5}$$

where e_i is i -th edge vector in the mesh and e'_i is the deformed version. Using the same principle in Section 3.1.1, we describe the scale factor s_e in terms of the coordinates of the unknown deformed edge vertices. Thus the edge distortion energy is defined as

$$E_e = \sum_{i=1}^M E_{eL}(i) + \sum_{i=1}^M E_{eR}(i) \tag{6}$$

where M is the number of edges in a frame. With this model, we smooth the edges of adjacent quads.

3.2 Disparity mapping energy

3.2.1 Spatial domain

The aim of the disparity mapping energy is to constrain the disparity change to ensure depth perception. We first design the disparity constraints in the spatial domain, to control the new disparity in both the horizontal and vertical directions. After energy-based optimization, we perform a 2D spatial projective transformation to every quad in the frames. This gives the new locations of the corresponding feature points using the vertex coordinates of the quads containing them. Thus, based on elementary geometry, we can describe the new disparity using a function f of the eight

unknown x -coordinates of the quads' vertices in the warped left and right frames. The detailed expression for the function f is omitted for the sake of brevity.

$$D_h(i) = f(x_{i,j}^L, x_{i+1,j}^L, x_{i+1,j+1}^L, x_{i,j+1}^L, x_{i,j}^R, x_{i+1,j}^R, x_{i+1,j+1}^R, x_{i,j+1}^R) \tag{7}$$

where x^L and x^R denote the vertices x -coordinates in the left and right frames respectively, i is the row index of the upper left vertex of a quad containing a feature point, and j is its column index. We now define the horizontal disparity mapping energy as

$$E_{dh} = \sum_{i=1}^K \|D'_h(i) - s_d D_h(i)\|^2 \tag{8}$$

where $D'_h(i)$ is the output horizontal disparity at the i -th pair of key points with original disparity $D_h(i)$, and K is the number of key point pairs. s_d is the scale factor, by which we can adjust the disparity linearly according to the viewing configuration. In Fig. 1 we increase disparity $1.5\times$. As mentioned in Ref. [3], vertical disparity always causes 3D fatigue for viewers. However, for points in the frames, the nonlinear transformation may cause location changes in the y -direction. Using the same principle, we can describe the vertical disparity and mapping energy as

$$D_v(i) = f(y_{i,j}^L, y_{i+1,j}^L, y_{i+1,j+1}^L, y_{i,j+1}^L, y_{i,j}^R, y_{i+1,j}^R, y_{i+1,j+1}^R, y_{i,j+1}^R) \tag{9}$$

$$E_{dv} = \sum_{i=1}^K \|D'_v(i)\|^2 \tag{10}$$

where $D'_v(i)$ is the output vertical disparity at the i -th pair of key points. Unlike in the case of the horizontal disparity, the objective is now to reduce the sum of the vertical disparity mapping energies to zero. Vertical disparity is one of the most important reasons for the dissatisfaction with 3D perception.

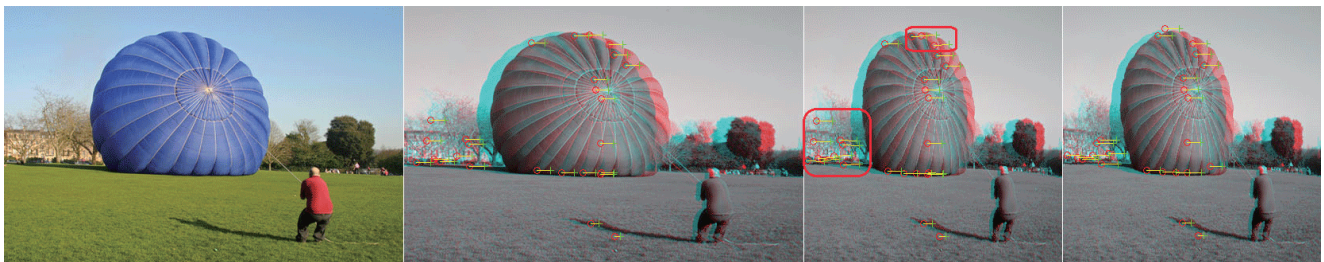


Fig. 2 Left to right: original left frame, original anaglyph frame, resizing without vertical disparity constraint, resizing with vertical disparity constraint.



Fig. 3 Resolution adaptation and depth remapping for the frame sequence. Left to right: original left frames (720×416), original disparity at the key points, left frames at target resolution (600×400), increased disparity ($1.5\times$).

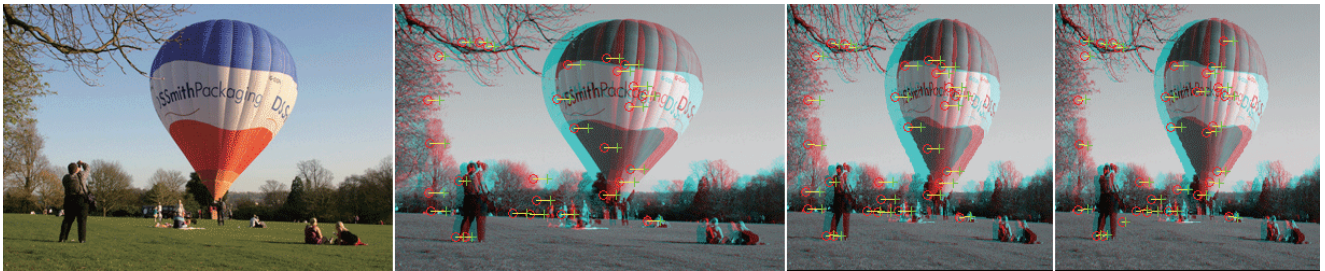


Fig. 4 Retargeted frame with and without spatial disparity control. Left to right: original left frames, original anaglyph frame, retargeted anaglyph frame with spatial disparity constraints, retargeted anaglyph frame without spatial disparity constraints.

See Fig. 2. With this constraint, vertical disparities are removed.

3.2.2 Temporal domain

We now consider how to constrain disparity consistency in the temporal domain. Many factors affect the detection of important contents in an image and cause different mesh deformation for the same objects in the adjacent frames. Performing retargeting on each frame separately will lead to depth jitter in certain local regions. Frequent disparity variation, i.e., depth jitter, easily causes visual fatigue even within the zone of comfortable viewing [5]. To resolve this problem, we collect the depth relations between each sequential pair of key frames by tracking the key points [30] and estimate their inter-frame disparity variation. In many cases, such as accelerating motion and so on, the disparity between adjacent key frames at a key point could be different. The disparity difference must be preserved

in the retargeted video. Also important is that the depth variation trend in the retargeted video must be consistent with that in the original video at the key points. To ensure that the depth changes are consistent, we preserve the disparity variation rate at every key point by defining a disparity jitter energy as

$$E_{dt} = \sum_{i=1}^Z \left\| D'_{T+1}(i) - \frac{D_{T+1}(i)}{D_T(i)} D'_T(i) \right\|^2 \quad (11)$$

where Z is the number of key points tracked successfully, $D'_T(i)$ is the output disparity value at time T at the i -th key point, and the original disparity value is $D_T(i)$. $D'_{T+1}(i)$ is the output disparity value at the same key point in the next key frame. (Frames are assumed to be one unit of time apart.) Note that at time $T + 1$, the original disparity variation rate and output disparity $D'_T(i)$ at time T are known constants.

Secondly, our concern is to preserve depth layer

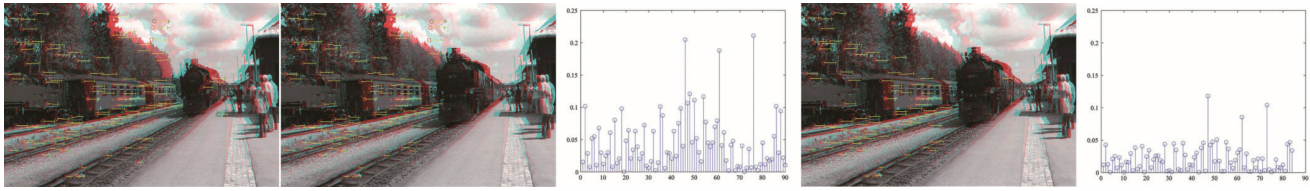


Fig. 5 Retargeted frame with and without temporal disparity control. Left to right: original anaglyph frame, retargeted anaglyph frame without temporal disparity constraints, difference in disparity variation rate without temporal disparity constraints, retargeted anaglyph frame with temporal disparity constraints, difference in disparity variation rate with temporal disparity constraints.

relations, i.e., intra-frame disparity variation. At a global level, any difference in depth between an object and its surroundings should be preserved in proportion. To simplify the problem, our goal is to keep the disparity differences at the key points changing in proportion to those between the original two neighboring frames. The constraint of inter-frame disparity variation can only ensure that the disparity differences between objects are unchanged. It cannot ensure that the proportional disparity differences change in adjacent frames. We define the number of the neighboring quads as the search window radius to find nearby key points. In our experiments, the number of neighboring quads is 2. We define the disparity difference energy as

$$E_{dd}(i) = \sum_{j=1}^{Z_j} \left\| \Delta D'_{T+1}(j) - \frac{\Delta D_{T+1}(j)}{\Delta D_T(j)} \Delta D'_T(j) \right\|^2 \quad (12)$$

$$E_{dd} = \sum_{i=1}^Z E_{dd}(i) \quad (13)$$

where Z_j is the number of nearby key points for the i -th key point. $\Delta D_T(j)$ is the disparity difference between the i -th key point and the j -th nearby key point at time T .

3.3 Energy optimization

The combined warp energy generated from all the aforementioned terms is

$$E = \omega_g E_g + \omega_e E_e + \omega_h E_{dh} + \omega_v E_{dv} + \omega_t E_{dt} + \omega_d E_{dd} \quad (14)$$

where ω_g , ω_e , ω_h , ω_v , ω_t , and ω_d are the weights of each energy component, which can be adjusted to suit needs.

The boundary constraint ensures that we cannot change the coordinates of the vertices at the four corners. The vertices on the boundaries are only allowed to move along the boundary edge. The minimization of the total energy constitutes a least-squares problem. By solving a system of linear

equations [33], we obtain the optimal results (Fig. 3). The weight values in Eq. (14), in order, are 1, 1, 5, 1000, 500, and 500, which denote that we pay more attention to removing the vertical disparity and the depth variation consistency in the temporal domain.

4 Experiments

Our method is efficient. Running on a PC with an Intel quad core 3.1 GHz CPU and 4 GB RAM, our method takes about 1.5 s for every stereoscopic frame at a resolution of 720×416 with quad size 20×20 . This is without any parallel algorithm based on either multicore CPU or GPU.

4.1 Without spatial disparity control

In Fig. 4, we retargeted one frame from resolution 362×248 to 248×248 twice. The spatial disparity constraints were applied in one case and not in the other. Note that \circ indicates a key point in the left frame and $+$ is the location of the corresponding point in the right frame. We could readily see the inconsistent disparity values and vertical disparities in the retargeted frame without the spatial disparity control.

4.2 Without temporal disparity control

In Fig. 5, we retargeted the sequence in Fig. 3 without the temporal disparity control. In the retargeting process, we tracked 90 key points in the sequence. Then we computed the disparity variation rate d_t/d_{t+1} for every corresponding point respectively and the differences of the rate values before and after retargeting. The smaller the difference value is, the better the disparity variation is preserved. The average difference value is 4.58%. In the same way, we obtained the average difference value is 2.19% when we considered the temporal disparity constraints. It shows that the disparity variation can be preserved better when we constrain the disparity in the temporal domain.



Fig. 6 Comparison of retargeting results. Top two images: two sequential original frames. Middle two images: results computed by the method in Ref. [21]. Bottom two images: results computed by our method. The objective is to reduce the disparity to lie within the comfort zone. In the blue circle, the disparity changes obviously when using Ref. [21]. Using our method, the disparity variation is more consistent with the original video.

4.3 Assessment

Lacking an objective assessment method for stereo video, and executable code for other methods, we cannot definitively state that our method is better than others. But by making a comparison to the video processing results given in previous work, we can demonstrate the effectiveness of our method. In Fig. 6, we compare the performance of our method to that of a previous video retargeting method [21]. The upper two images are two sequential original frames. The objective is to improve the visual experience by reducing excessive disparity. Figure 6 shows red-cyan pictures, clearly indicating the disparity range. The method in Ref. [21] does not take account of temporal disparity control and causes disparity variation for the white car, while the disparity is stable in the original video. However, by using our method, disparity variation is constrained. In the

two frames, the disparity values in the blue circle are reduced and kept similar.

We further assessed the perceptual quality of our method by performing a subjective assessment based on a user study. We adopted the pair comparison method with 4 videos from “Mobile 3DTV” and produced 3 test videos for each one [34, 35]. In short, as judged by 14 observers with normal stereopsis, we obtained good results: 86% correct discrimination for depth change and 81% acceptance of the picture quality.

4.4 Limitations

Via experimentation and subjective assessment, we may summarize the limitations of our method. Firstly, for videos containing many structured objects in the pictures, our method cannot produce good results because there is not enough space left for warping, which is a common problem

for warping-based methods. Secondly, too large a disparity adjustment will impair the quality of the video. Layer-based resolution adaptation methods or modification based on monocular depth cues might be helpful in overcoming this problem in some cases.

5 Conclusions

In this paper, we proposed a retargeting method to resize stereoscopic 3D video and adjust disparity of video content to lie in the comfort zone. We build energy models for picture deformation and disparity mapping. To preserve important image contents, we constrain grid deformation and edge distortion. To maintain depths within the comfort zone, we constrain the disparity in the spatial and temporal domains. The spatial disparity constraints apply to horizontal and vertical disparity mapping. The temporal disparity constraints consider inter-frame disparity variation and intra-frame disparity variation. Finally, by adjusting the weight combination, our method can be applied to various view configurations. The objective of our work is to perform content-aware retargeting of pictures of video frames and meanwhile adjust the disparity values for the contents in stereoscopic video. Although there are some limitations to our method, our experiments have demonstrated its capabilities by a user study and a comparison to the latest related work. We hope that our work promotes future research in stereoscopic 3D media retargeting.

Acknowledgements

This work was supported by the National Basic Research Program of China under Grant No. 2011CB302206, the National Natural Science Foundation of China under Grant Nos. 61272226 and 61272231, and Beijing Key Laboratory of Networked Multimedia.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

[1] Shamir, A.; Sorkine, O. Visual media retargeting. In:

- ACM SIGGRAPH ASIA 2009 Courses, Article No. 11, 2009.
- [2] Koppal, S. J.; Zitnick, C. L.; Cohen, M.; Kang, S. B.; Ressler, B.; Colburn, A. A viewer-centric editor for 3D movies. *IEEE Computer Graphics and Applications* Vol. 31, No. 1, 20–35, 2011.
- [3] Mendiburu, B. *3D Movie Making: Stereoscopic Digital Cinema from Script to Screen*. CRC Press, 2012.
- [4] Kooi, F. L.; Toet, A. Visual comfort of binocular and 3D displays. *Displays* Vol. 25, Nos. 2–3, 99–108, 2004.
- [5] Lambouij, M.; Fortuin, M.; Heynderickx, I.; IJsselstein, W. Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology* Vol. 53, No. 3, 30201, 2009.
- [6] Chen, L.-Q.; Xie, X.; Fan, X.; Ma, W.-Y.; Zhang, H.-J.; Zhou, H.-Q. A visual attention model for adapting images on small displays. *Multimedia Systems* Vol. 9, No. 4, 353–364, 2003.
- [7] Avidan, S.; Shamir, A. Seam carving for contentaware image resizing. *ACM Transactions on Graphics* Vol. 26, No. 3. Article No. 10, 2007.
- [8] Rubinstein, M.; Shamir, A.; Avidan, S. Improved seam carving for video retargeting. *ACM Transactions on Graphics* Vol. 27, No. 3, Article No. 16, 2008.
- [9] Gal, R.; Sorkine, O.; Cohen-Or, D. Feature-aware texturing. In: Proceedings of the 17th Eurographics conference on Rendering Techniques, 297–303, 2006.
- [10] Wolf, L.; Guttmann, M.; Cohen-Or, D. Nonhomogeneous content-driven video-retargeting. In: IEEE 11th International Conference on Computer Vision, 1–6, 2007.
- [11] Wang, Y.-S.; Tai, C.-L.; Sorkine, O.; Lee, T.-Y. Optimized scale-and-stretch for image resizing. *ACM Transactions on Graphics* Vol. 27, No. 5, Article No. 118, 2008.
- [12] Lin, H.-Y.; Chang, C.-C.; Hsieh, C.-H. Cooperative resizing technique for stereo image pairs. In: 2012 International Conference on Software and Computer Applications, Vol. 41, 96–100, 2012.
- [13] Basha, T. D.; Moses, Y.; Avidan, S. Stereo seam carving a geometrically consistent approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 35, No. 10, 2513–2525, 2013.
- [14] Chang, C.-H.; Liang, C.-K.; Chuang, Y.-Y. Contentaware display adaptation and interactive editing for stereoscopic images. *IEEE Transactions on Multimedia* Vol. 13, No. 4, 589–601, 2011.
- [15] Niu, Y.; Feng, W.-C.; Liu, F. Enabling warping on stereoscopic images. *ACM Transactions on Graphics* Vol. 31, No. 6, Article No. 183, 2012.
- [16] Luo, S.-J.; Shen, I.-C.; Chen, B.-Y.; Cheng, W.-H.; Chuang, Y.-Y. Perspective-aware warping for seamless stereoscopic image cloning. *ACM Transactions on Graphics* Vol. 31, No. 6, Article No. 182, 2012.
- [17] Du, S.; Hu, S.; Martin, R. Changing perspective in stereoscopic images. *IEEE Transactions on Visualization and Computer Graphics* Vol. 19, No. 8, 1288–1297, 2013.

- [18] Tong, R.-F.; Zhang, Y. Cheng, K.-L. StereoPasting: Interactive composition in stereoscopic images. *IEEE Transactions on Visualization and Computer Graphics* Vol. 19, No. 8, 1375–1385, 2013.
- [19] Lin, H.-S.; Guan, S.-H.; Lee, C.-T.; Ouhyoung, M. Stereoscopic 3D experience optimization using cropping and warping. In: SIGGRAPH Asia 2011 Sketches, Article No. 40, 2011.
- [20] Lee, K.-Y.; Chung, C.-D.; Chuang, Y.-Y. Scene warping: Layer-based stereoscopic image resizing. In: IEEE Conference on Computer Vision and Pattern Recognition, 49–56, 2012.
- [21] Lang, M.; Hornung, A.; Wang, O.; Poulakos, S.; Smolic, A.; Gross, M. Nonlinear disparity mapping for stereoscopic 3D. *ACM Transactions on Graphics* Vol. 29, No. 4, Article No. 75, 2010.
- [22] Yan, T.; Lau, R. W. H.; Xu, Y.; Huang, L. Depth mapping for stereoscopic videos. *International Journal of Computer Vision* Vol. 102, Nos. 1–3, 293–307, 2013.
- [23] Chamaret, C.; Boisson, G.; Chevance, C. Video retargeting for stereoscopic content under 3D viewing constraints. In: Proceedings of SPIE 8288, Stereoscopic Displays and Applications XXIII, 82880H, 2012.
- [24] Liu, C.-W.; Huang, T.-H.; Chang, M.-H.; Lee, K.-Y.; Liang, C.-K.; Chuang, Y.-Y. 3D cinematography principles and their applications to stereoscopic media processing. In: Proceedings of the 19th ACM international conference on Multimedia, 253–262, 2011.
- [25] Templin, K.; Didyk, P.; Myszkowski, K.; Hefeida M. M.; Seidel H.-P.; Matusik, W. Modeling and optimizing eye vergence response to stereoscopic cuts. *ACM Transactions on Graphics* Vol. 33, No. 4, Article No. 145, 2014.
- [26] Mu, T.-J.; Sun, J.-J.; Martin, R. R.; Hu, S.-M. A response time model for abrupt changes in binocular disparity. *The Visual Computer* Vol. 31, No. 5, 675–687, 2014.
- [27] Harel, J.; Koch, C.; Perona, P. Graph-based visual saliency. In: Advances in Neural Information Processing Systems 19, 545–552, 2006.
- [28] Bay, H.; Tuytelaars, T.; Gool, L. V. SURF: Speeded up robust features. *Lecture Notes in Computer Science* Vol. 3951, 404–417, 2006.
- [29] Fischler, M. A.; Bolles, R. C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* Vol. 24, No. 6, 381–395, 1981.
- [30] Shi, J.; Tomasi, C. Good features to track. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 593–600, 1994.
- [31] Von Gioi, R. G.; Jakubowicz, J.; Morel, J.-M.; Randall, G. LSD: A line segment detector. *Image Processing On Line* Vol. 2, 35–55, 2012.
- [32] Krähenbühl, P.; Lang, M.; Hornung, A.; Gross, M. A system for retargeting of streaming video. *ACM Transactions on Graphics* Vol. 28, No. 5, Article No. 126, 2009.
- [33] Paige, C. C.; Saunders, M. A. LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software* Vol. 8, No. 1, 43–71, 1982.
- [34] Smolic, A.; Tech, G.; Brust, H. Report on generation of stereo video data base. Mobile3DTV Technical Report, 2010.
- [35] ITU-T RECOMMENDATION P. Subjective video quality assessment methods for multimedia applications. 1999. Available at <https://www.itu.int/rec/T-REC-P.910/en>.



Yi Liu received the B.S. and M.S. degrees from China Ordnance Engineering College, Shijiazhuang, China. Since 2005, he has been a researcher at the Beijing Aerospace Control Center, Beijing, China. He is currently working toward the Ph.D. degree at the Department of Computer

Science and Technology, Tsinghua University, Tsinghua NLIST, Beijing, China. His research interests include 3D image/video processing, computer vision, human depth perception, and visual quality assessment.



Lifeng Sun received the B.S., M.S., and Ph.D. degrees in systems engineering from the National University of Defense Technology, Changsha, China, in 1995, 1997, and 2000, respectively. He is currently a professor in the Department of Computer Science and Technology, Tsinghua University, Beijing, China.

His research interests include interactive multiview video, video sensor networks, peer-to-peer streaming, and distributed video coding.



Shiqiang Yang graduated from the Department of Computer Science and Technology, Tsinghua University, Beijing, China, in 1977, and received the M.E. degree in 1983. He is now a professor at Tsinghua University. His research interests include multimedia technology and systems, video compression and streaming, content-based retrieval, semantics for multimedia information, and embedded multimedia systems. He has published more than 100 papers in international conferences and journals. He is currently the president of the Multimedia Committee of the China Computer Federation.

Other papers from this open access journal are available at no cost from <http://www.springer.com/journal/41095>. To submit a manuscript, please go to <https://www.editorialmanager.com/cvmj>.