

Optimizing visual comfort for stereoscopic 3D display based on color-plus-depth signals

Feng Shao,* Qiuping Jiang, Randi Fu, Mei Yu, and Gangyi Jiang

Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China
*shaofeng@nbu.edu.cn

Abstract: Visual comfort is a long-facing problem in stereoscopic 3D (S3D) display. In this paper, targeting to produce S3D content based on color-plus-depth signals, a general framework for depth mapping to optimize visual comfort for S3D display is proposed. The main motivation of this work is to remap the depth range of color-plus-depth signals to a new depth range that is suitable to comfortable S3D display. Towards this end, we first remap the depth range globally based on the adjusted zero disparity plane, and then present a two-stage global and local depth optimization solution to solve the visual comfort problem. The remapped depth map is used to generate the S3D output. We demonstrate the power of our approach on perceptually uncomfortable and comfortable stereoscopic images.

©2016 Optical Society of America

OCIS codes: (330.1400) Vision - binocular and stereopsis; (120.2040) Displays.

References and links

1. H. J. Yeom, H. J. Kim, S. B. Kim, H. Zhang, B. Li, Y. M. Ji, S. H. Kim, and J. H. Park, "3D holographic head mounted display using holographic optical elements with astigmatism aberration compensation," *Opt. Express* **23**(25), 32025–32034 (2015).
2. D. Zhao, B. Su, G. Chen, and H. Liao, "360 degree viewable floating autostereoscopic display using integral photography and multiple semitransparent mirrors," *Opt. Express* **23**(8), 9812–9823 (2015).
3. C. Lee, G. Seo, J. Lee, T. H. Han, and J. G. Park, "Auto-stereoscopic 3D displays with reduced crosstalk," *Opt. Express* **19**(24), 24762–24774 (2011).
4. N. S. Holliman, N. A. Dodgson, G. E. Favalora, and L. Pockett, "Three-dimensional displays: A review and applications analysis[J]. Broadcasting," *IEEE Trans. Broadcast* **57**(2), 362–371 (2011).
5. D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue," *J. Vis.* **8**(3), 1–30 (2008).
6. Y. Mori, N. Fukushima, T. Yendo, T. Fujii, and M. Tanimoto, "View generation with 3D warping using depth information for FTV," *Signal Process. Image Commun.* **24**(1–2), 65–72 (2009).
7. K. H. Yoon, H. Ju, I. Park, and S. K. Kim, "Determination of the optimum viewing distance for a multi-view auto-stereoscopic 3D display," *Opt. Express* **22**(19), 22616–22631 (2014).
8. M. Urvoy, M. Barkowsky, and P. Le Callet, "How visual fatigue and discomfort impact 3D-TV quality of experience: A comprehensive review of technological, psychophysical, and psychological factors," *Ann. Telecommun.* **68**(11–12), 641–655 (2013).
9. M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and M. Gross, "Nonlinear disparity mapping for stereoscopic 3D," *ACM Trans. Graph.* **29**(4), 75 (2010).
10. P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski, and H.-P. Seidel, "E. Eisemann E, H. P. Seidel, and W. Matusik, "A perceptual model for disparity," *ACM Trans. Graph.* **30**(4), 96 (2011).
11. P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski, H. P. Seidel, and W. Matusik, "A luminance-contrast-aware disparity model and applications," *ACM Trans. Graph.* **31**(6), 184 (2012).
12. A. Chapiro, S. Heinzle, T. O. Aydın, S. Poulakos, M. Zwicker, A. Smolic, and M. Gross, "Optimizing stereo-to-multiview conversion for autostereoscopic displays," *Proc. of Computer Graphics Forum* **33**(2), 63–72 (2014).
13. J. Lei, S. Li, B. Wang, K. Fang, and C. Hou, "Stereoscopic visual attention guided disparity control for multiview images," *J. Disp. Technol.* **10**(5), 373–379 (2014).
14. F. Shao, Z. Li, Q. Jiang, G. Jiang, M. Yu, and Z. Peng, "Visual discomfort relaxation for stereoscopic 3D images by adjusting zero-disparity plane for projection," *Displays* **39**, 125–132 (2015).
15. H. Sohn, Y. J. Jung, S. Lee, and F. Speranza, "Visual comfort amelioration technique for stereoscopic images: Disparity remapping to mitigate global and local discomfort causes," *IEEE Trans. Circ. Syst. Video Tech.* **24**(5), 745–758 (2014).
16. Y. J. Jung, H. Sohn, S. Lee, and Y. M. Ro, "Visual comfort improvement in stereoscopic 3D displays using perceptually plausible assessment metric of visual comfort," *IEEE Trans. Consum. Electron.* **60**(1), 1–9 (2014).

17. C. Oh, B. Ham, S. Choi, and K. Sohn, "Visual fatigue relaxation for stereoscopic video via nonlinear disparity remapping," *IEEE Trans. Broadcast* **6**(2), 142–153 (2015).
18. C. Jung, L. Cao, H. Liu, and J. Kim, "Visual comfort enhancement in stereoscopic 3D images using saliency-adaptive nonlinear disparity mapping," *Displays* **40**, 17–23 (2015).
19. T. Yan, R. W. H. Lau, Y. Xu, and L. Huang, "Depth mapping for stereoscopic videos," *Int. J. Comput. Vis.* **102**(1–3), 293–307 (2013).
20. D. Kim, S. Choi, and K. Sohn, "Depth adjustment for stereoscopic images and subjective preference evaluation," *J. Electron. Imaging* **20**(3), 033011 (2011).
21. G. Damberg and W. Heidrich, "Efficient freeform lens optimization for computational caustic displays," *Opt. Express* **23**(8), 10224–10232 (2015).
22. S. W. Jung and S. J. Ko, "Depth enhancement considering just noticeable difference in depth," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* **95**(3), 673–675 (2012).
23. S. W. Jung and S. J. Ko, "Depth sensation enhancement using the just noticeable depth difference," *IEEE Trans. Image Process.* **21**(8), 3624–3637 (2012).
24. S. W. Jung, "A modified model of the just noticeable depth difference and its application to depth sensation enhancement," *IEEE Trans. Image Process.* **22**(10), 3892–3903 (2013).
25. J. Lei, C. Zhang, Y. Fang, Z. Gu, N. Ling, and C. Hou, "Depth sensation enhancement for multiple virtual view rendering," *IEEE Trans. Multimed.* **17**(4), 457–469 (2015).
26. T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks, "The zone of comfort: Predicting visual discomfort with stereo displays," *J. Vis.* **11**(8), 11 (2011).
27. S. P. Du, B. Masia, S. M. Hu, and D. Gutierrez, "A metric of visual comfort for stereoscopic motion," *ACM Trans. Graph.* **32**(6), 222 (2013).
28. Q. Jiang, F. Shao, W. Lin, G. Jiang, and M. Yu, "On predicting visual comfort of stereoscopic images: A learning to rank based approach," *IEEE Signal Process. Lett.* **23**(2), 302–306 (2016).
29. D. V. S. X. De Silva, E. Ekmekcioglu, W. A. C. Fernando, and S. T. Worrall, "Display dependent preprocessing of depth maps based on just noticeable depth difference modeling," *IEEE J. Sel. Top. Signal Process.* **5**(2), 335–351 (2011).
30. 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.* **38**, 57–69 (2015).
31. ITU-R BT-500.11, "Methodology for the subjective assessment of the quality of television pictures," ITU-R BT-500.11, 2002.
32. ITU-R BT.1438, "Subjective assessment for stereoscopic television pictures," ITU-R BT.1438 (2000).
33. ITU-T P.910, "Subjective video quality assessment methods for multimedia applications," Recommendation ITU-T P.910, ITU Telecom. Sector of ITU (1999).
34. ITU-T P.911, Subjective video quality assessment methods for multimedia applications, Recommendation ITU-T P.911, ITU Telecom. Sector of ITU (1999).

1. Introduction

Stereoscopic 3D (S3D) contents have become very popular in recent years [1,2]. Especially, after the milestone success of "Avatar", more and more 3D movies are being produced in 3D, which also accelerate the popularization of 3D displays. Currently, many 3D representation formats had been proposed for display, including stereoscopic images, multi-view video plus depth (MVD), color-plus-depth (V + D), etc. However, in contrast to these standard 3D content formats, 3D display still lacks a definite standard for glass-based or glass-free displays [3,4]. Comparatively speaking, 2-view S3D display is just maturing. Consequently, S3D content production through format conversion will likely be a feasible way for ensuring backward compatible with S3D display.

Viewing a scene through a S3D display differs from natural viewing - it introduces accommodation-vergence (AV) conflict [5]. The main technical challenge in the 3D content production is that the disparity range of S3D content will exceeds the limitation of comfort zone, while current production methods, such as depth-image based rendering (DIBR) [6], directly synthesizes two views, and do not take the inherent depth range limitations into account. Moreover, if we watch a 3D movie at different viewing distances, the depth ranges for the comfort zone should be not the same [7], e.g., smaller depth range for small viewing distance, and vice versa. Thus, we expect any S3D content production should not only do not reduce the perceived visual comfort, but also adapt to different display settings [8]. Previous works solve the challenges through several ways:

- 1) *Disparity manipulation* technique has been proposed to adjust S3D images based on various manipulation operators. Existing disparity manipulation technologies [9–12] directly edit the original disparity of S3D contents to meet the requirements.

However, such approaches require manipulation operators for guidance, and more importantly, they do not target visual comfort enhancement.

- 2) *Disparity shifting* is a useful technique for improving visual comfort, which adjusts the zero disparity plane (ZDP) of a scene while maintaining its original disparity range [13,14]. Through disparity shifting, excessive disparities can be reduced to minimize the AV conflicts. However, disparity shifting alone cannot fully cover the comfort zone range if the disparity range exceeds a certain comfortable viewing zone. As a result, the usability of the technique is restricted.
- 3) *Disparity remapping* aims to (linearly or non-linearly) scale the disparity range into a comfort zone to avoid excessive AV conflicts [15–20]. However, since the comfort zone range is limited, if large disparity range scaling is conducted, rendering artifacts will be appeared in the processed images. Overall, disparity remapping may be a promising tool for improving visual comfort.

In this paper, inspired by the disparity shifting and disparity remapping frameworks, we provide a solution to solve the visual discomfort problem. The challenge of this problem is to build a bridge between S3D visual comfort and depth sensation based on color-plus-depth signals. We build a two-stage global and local optimization solution based on the visual comfort prediction function and just noticeable depth difference (JNDD) model. The output S3D content is generated by applying depth mapping to the original color-plus-depth signals using a warping-based technology. To the best of our knowledge, the joint visual comfort and depth sensation optimization problem for color-plus-depth signals has not been well addressed. In summary, our paper makes the following contributions:

- We provide a visual comfort optimization solution for color-plus-depth signals based on two complementary processes: global and local depth mapping. They are designed to optimize visual comfort and depth sensation simultaneously for S3D display.
- By taking the depth confidence and depth difference constraints into consideration, a global depth mapping optimization is designed to remap the depth ranges of all depth planes to assure globally comfortable viewing.
- By further iteratively adjusting the depth range of each depth plane, the globally remapped depths are locally enhanced for each plane to satisfy the visual comfort and depth differences constraints.
- Comprehensive objective and subjective validations of the approach are conducted on comfortable and uncomfortable stereoscopic images, demonstrating its general quality optimization performance.

In the remainder of this paper, we review the related work in Section 2, detail our method in Section 3, and finally present results in Section 4 and discussion in Section 5.

2. Related works

The area of S3D display has been researched extensively in the last decade, and many works have been devoted to improve S3D display design, leading to a recent trend towards computational displays [21]. However, content production for S3D display still poses an unresolved challenge. Recently, “color-plus-depth” format has emerged as a promising representation for 3D scene, to generate multiview/virtual images via DIBR or image warping techniques. However, these approaches do not take the inherent visual comfort limitations into account.

Since our main goal is to implement depth/disparity manipulations of S3D content, we will give an overview of existing depth/disparity adjustment methods. The most direct way is to work on pixel’s disparity to map the scene depth into a comfort zone. Lang *et al.* [9] presented a solution for general nonlinear disparity mapping operators for stereoscopic 3D, that uses disparity and saliency estimates to compute a deformation of the input views to meet

the target disparities. Didyk *et al.* [10] proposed a perceptual disparity model to predict human disparity perception by identifying the interaction of disparity magnitude and spatial frequency. Didyk *et al.* [11] presented a model to capture the interaction of disparity and luminance contrast, which can automatically retarget disparity and manipulate luminance contrast to improve depth perception. Chapiro *et al.* [12] presented a saliency-based stereo-to-multiview conversion method that generates optimized content for autostereoscopic multiview displays. However, these methods aim to enhance the visual experience for well-produced S3D content, and their potential applications in visual comfort enhancement are still unknown.

The motivation of visual comfort enhancement is to fit the disparity range into a limited depth range (called comfort zone) where the conflict between accommodation and vergence is reduced. Towards this end, state-of-the-art visual comfort enhancement methods follow two brand categories: disparity shifting [13,14] and disparity remapping [15–20]. The disparity shifting methods is comparatively simply that adjust the zero disparity plane (ZDP) of a scene but maintain its overall disparity range. Towards this end, visual attention model and visual discomfort map are guided for key-point selection in [13] and [14], respectively. However, disparity shifting alone cannot cover the comfortable viewing zone if the disparity range exceeds the range of a certain comfortable viewing zone. Linear disparity remapping is the simplest way to scale a disparity range to the comfortable viewing zone (CVZ). However, unnatural visual artifacts may occur by the linear way. Recently, nonlinear disparity remapping schemes have been presented to enhance 3D viewing experience. Nonlinear disparity mapping operator is a practical technique in content postproduction. In Sohn *et al.*'s series of works, various visual comfort improvement methods are proposed based on linear and nonlinear disparity remapping processes jointly to mitigate global and local discomfort causes in [15], by adjusting (scaled and shifted) the disparity of a scene under the guidance of an objective VCA metric in [16], and based on a nonlinear remapping operator to adjust the disparity range according to the determined visual fatigue score in [17]. Jung *et al.* [18] proposed a saliency-adaptive nonlinear disparity mapping based on a sigmoid function to minimize disparity distortions. Other relevant works can also be found in [19,20]. However, these visual comfort enhancement models do not consider the influence of RGB image content on depth perception.

From another perspective, depth sensation enhancement plays an important role in stereoscopic 3D experience. Many depth sensation enhancement methods have been proposed [22–25]. The most important depth sensation methods are designed based on JNDD model. The JNDD threshold is the smallest change in a depth stimulus that human cannot perceive the change. Those JNDD based depth sensation methods aim to increase the depth difference between neighboring objects. In Jung *et al.*'s series of works, a global depth sensation enhancement algorithm is developed in [22], which stretch the depth differences between layers via energy minimization. In [23], an improved JNDD-based depth sensation enhancement algorithm is proposed via an energy minimization framework using three energy terms defined as depth data preservation, depth-order preservation, and depth difference expansion. In [24], a modified JNDD measurement method is proposed that adjusts the physical size of the object such that the perceived size of the object is maintained. In Lei *et al.*'s work [25], depth sensation enhancement is applied to the semantic of virtual view based on an energy function built by the number of rendered views and the saliency analysis.

To our knowledge, although the existing approaches can address the visual comfort and depth sensation optimization issues to a certain extent, how to optimize them simultaneously and how to apply them to color-plus-depth signals for S3D display, still remain not fully investigated. To address the above issues, we try to build a bridge between visual comfort and depth sensation under different 3D viewing conditions. For this purpose, we design a new depth mapping framework to simultaneously optimize visual comfort and depth sensation for S3D display, and subjective perception experiments are conducted to validate the framework.

3. Proposed approach

Our method aims to generate visually comfortable S3D output from color-plus-depth signals suitable for S3D display. The overall pipeline is illustrated in Fig. 1. First, by determining the crossed and uncrossed disparity ranges before and after ZDP adjustment, the depth range is globally remapped to a new depth range. Then, a two-stage global and local depth mapping optimization is conducted. The global optimization is designed to assure globally comfortable viewing, while the following local optimization is further to assure locally enhanced visual comfort. Finally, the remapped depth map is used to generate the final stereopair output. In the following, we will give more details on the individual steps.

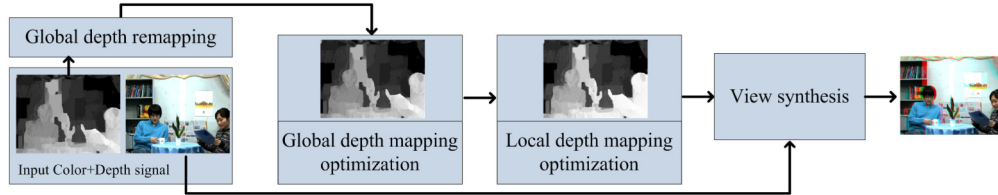


Fig. 1. Illustration of our proposed depth mapping model.

3.1. Global depth remapping

Since the depth information in a color-plus-depth signal can be generated by various means of depth estimation, depth capturing or manual compilation, the initial depth range is usually not well suitable for S3D generation, which tends to support a significantly discomfort depth range. The main motivation of this work is to remap the original depth range of a color-plus-depth signal to a new depth range by changing the depths. Let Z and \hat{Z} be the perceived depths by the viewer when watching a S3D signal on a target screen before and after depth remapping process, respectively. We also denote $[Z_{\min}, Z_{\max}]$ and $[\hat{Z}_{\min}, \hat{Z}_{\max}]$ be the original and remapped depth ranges, respectively. Hence, the original depth Z can be linearly (also globally or locally) remapped to a new depth \hat{Z} to avoid excessive AV conflicts of the scene as follows:

$$\hat{Z} = k(Z - Z_{\min}) + \hat{Z}_{\min} \quad (1)$$

where

$$k = \frac{\hat{Z}_{\max} - \hat{Z}_{\min}}{Z_{\max} - Z_{\min}} \quad (2)$$

Researchers working on visual discomfort have found that the range of comfort zone is limited, e.g., $\pm 1^\circ$ disparity for the Percival's Zone of Comfort (PZC) [26]. Of course, comfortable depth range is also related to the 3D content, such as motion and disparity changing [27]. Thus, if we directly squeeze large depth range of 3D signals to their favorite depth range, rendering artifacts will be appeared in the synthesized images.

Let V be the distance between the viewer and the screen, e be the human interpupillary distance, and s be the corresponding disparity for a pixel of a stereo image, the depth perceived by the viewer can be determined as follows:

$$Z = \frac{e \cdot V}{e - s} \quad (3)$$

From (3), it is clear that the minimum and maximum depths are determined by the minimum and maximum disparity range limits (e.g., the crossed and uncrossed disparity ranges). A direct way is to adjust the crossed and uncrossed disparity ranges. Based on our

previous work [14] that adjusts the ranges of the crossed and uncrossed disparities, we first select an optimal ZDP guided by the predicted visual discomfort map.

Then, taking the ZDP as convergence plane, the perceived depth is re-defined as

$$\hat{Z} = \frac{e \cdot V}{e - (s - s_{ZDP})} = \frac{1}{\left(\frac{1}{Z_{ZDP}} + \frac{1}{Z}\right) - \frac{1}{V}} \quad (4)$$

Here, $s_{ZDP} = e - eV / Z_{ZDP}$ is the disparity for the convergence plane based on the depth value Z_{ZDP} in the ZDP. Then, the minimum and maximum depths before and after ZDP adjustment can be easily calculated. Finally, the depth range is remapped using Eq. (1). The calculation of the ZDP can refer to our previous paper [14] for details.

Since the overall disparity range after adjustment is not affected, the overall depth range is also not largely changed but placed into a comparatively comfortable depth scale. However, since all depth planes are remapped using a same scale factor, it may lead to some cases, e.g., comfortable planes become more comfortable, or comfortable planes become uncomfortable. In order to optimize the process, we further implement global and local mapping optimization on each depth plane.

3.2. Global depth mapping optimization

Since we aim to control the depth range of color-plus-depth signals to enhance the visual comfort but preserve the depth sensation, different depth layers are regarded as basic units in this study. In order to divide a depth map into multiple depth planes, we first identify the depth bins from the depth histogram, and determine the number of depth planes. For example, if four local minima are identified in the histogram, the depth map will be divided into three planes. With the determined number of planes, a simple k-means clustering algorithm is performed on the depth map to derive the plane information.

Our pipeline starts by globally transforming the input pixel-wise depth space into a new plane-wise depth space that better matches the comfort zone limit for stereoscopic production. In order to ensure global visual comfort, the adjustment of depth range for all planes is synchronized. Let p_i^o represent the average depth value of the i -th depth plane of the original depth map, the target average depth values for all planes can be found by minimizing the following energy function:

$$\begin{aligned} p^* &= \arg \min_p \{E_{DATA}(p^o, p) + E_{JNDD}(p^o, p)\} \\ \text{s.t. } \Phi(\mathbf{f}) &> VC_{GT} \end{aligned} \quad (5)$$

where $\Phi(\cdot)$ is a pre-defined visual comfort prediction function, and the threshold VC_{GT} is set to 4 to match the ‘comfortable’ scale in mean opinion score (MOS) value. The global feature vector \mathbf{f} is computed using the same method in the visual comfort prediction function. p^o and p^* are the sets of average depth values of the original and adjusted depth planes with the forms as follows:

$$p^o = \{p_1^o, p_2^o, \dots, p_L^o\}, \quad p^* = \{p_1^*, p_2^*, \dots, p_L^*\}. \quad (6)$$

In order to predict the degree of visual comfort of stereoscopic images, we learn a visual comfort prediction function from a given training database. Refer to our previous work [28], the same fifty samples (10 samples for each visual comfort scale) that cover all possible ranges of visual comfort scale are selected from our NBU S3D-VCA database. Then, given the factors (e.g., excessive binocular disparity, defocus blur and low spatial frequency of 3D stimulus) that are more likely to evoke visual discomfort, we extract the visual comfort-aware features for each sample in accordance with these three influential factors, respectively. Considering the absolute scales used in absolute categorical rating (ACR) test as different

rank values of visual comfort, we learn a ranking model for VCA, denoted as $\Phi(\cdot)$. The learned ranking model can be used to predict a ranking score for an arbitrary stereoscopic image to indicate the degree of visual discomfort present in it. The learned model has higher generalization capability in predicting the visual comfort. Detailed information can be found in our previous work [28].

Recall that the depth range in Eq. (1) is defined to remap from $[Z_{\min}, Z_{\max}]$ to $[\hat{Z}_{\min}, \hat{Z}_{\max}]$. In our global depth mapping optimization, we linearly remap all depth planes into their target planes. If we define $R = Z_{\max}^i - Z_{\min}^i$ and $R' = \hat{Z}_{\max}^i - \hat{Z}_{\min}^i$ for each plane, the linear function $\varphi_i: [Z_{\min}^i, Z_{\max}^i] \rightarrow [\hat{Z}_{\min}^i, \hat{Z}_{\max}^i]$ is defined as the form:

$$\varphi_i(x) = \gamma_i x + \alpha_i \quad (7)$$

where $\gamma_i = R'/R$ to control the scale, and α_i is a compensation factor. Inversely speaking, the defined linear function is to adjust the depth range of each plane. The first term $E_{DATA}(p^o, p)$ is used to enforce the minimal variety of depth values in the depth map, and it is defined as

$$E_{DATA}(p^o, p) = \sum_{i=1}^L \beta_i (p_i - p_i^o) \quad (8)$$

where β_i controls the depth confidence of each plane.

The second term $E_{JNDD}(p^o, p)$ enables the stretching of depth difference, and is defined as follows:

$$E_{JNDD} = \lambda_1 \cdot \sum_{i=1}^L \sum_{j \in \Omega_i} \max(0, D_{JNDD}(p_i^o) - |p_i - p_j|) \quad (9)$$

where Ω_i denotes a set of indices of the planes connected to the i -th plane; λ_1 is a weighting factor. This term is effective only when the recorded depth difference between the connected planes is less than D_{JNDD} .

The relationship between the JNDD (the minimal difference between objects that can be perceived by human's eyes) and depth level has been modeled through subjective evaluation. Specifically, the numerical formula for the JNDD is given as [29]

$$D_{JNDD}(d) = \begin{cases} 21, & \text{if } 0 \leq d \leq 64 \\ 19, & \text{if } 64 \leq d \leq 128 \\ 18, & \text{if } 128 \leq d \leq 192 \\ 20, & \text{if } 192 \leq d \leq 255 \end{cases} \quad (10)$$

where d represents the pixel-wise depth value of a depth map. Generally, the JNDD model is related to the types of display devices, and thus it should be refined with different display environments [23].

Given the p^o and p^* , if the original depth values satisfying $p_i^o \leq d^o(x, y) \leq p_{i+1}^o$, the enhanced depth value $d^*(x, y)$ is obtained by

$$d^*(x, y) = \frac{p_{i+1}^* - p_i^*}{p_{i+1}^o - p_i^o} (d^o(x, y) - p_i^o) + p_i^* \quad (11)$$

Refer to [23], Eq. (5) is solved by the genetic algorithm, which is a powerful tool to solve nonlinear and nonconvex objective functions. The solution of (5) is also dependent on the

factors β_i and λ_i . In our experiments, β_i is fixed as $\{0.5, 4, 2\}$ for $i = \{1, 2, 3\}$, and $\omega_1 = 100$. Since the number of objects is not fixed, these parameters will be changed accordingly.

3.3. Local depth mapping optimization

After the above global depth mapping optimization, we perform an additional local depth mapping optimization process. The main goal of this step is to locally enhance the visual comfort for those important regions. Our previous work [30] has been revealed that strong depth sensation comes from those extremely uncomfortable regions under uncomfortable 3D viewing, while those extremely comfortable regions are usually accompanied by lower depth sensation under comfortable 3D viewing. From this perspective, another local depth mapping optimization for each depth plane is necessary to compensate these factors. Different from the existing depth sensation methods [22–25] that only take JNDD constraint into account, we simultaneously take the visual comfort and depth sensation constraints into account for optimization.

Let \mathbf{f}_i^* be the feature vector for the i -th depth plane (computing by the same method with global feature vector \mathbf{f}), the resultant enhanced average depth value is defined as

$$\begin{aligned} p_i^* &\approx \gamma_i \cdot p_i^o + \alpha_i \\ \text{s.t. } &\begin{cases} \Phi(\mathbf{f}_i^*) \geq VC_{GT} \\ \forall j, |p_i^* - p_j^o| \geq \max_{j \in \Omega} \{D_{JNDD}(p_j^o)\} \end{cases} \end{aligned} \quad (12)$$

The above process is optimized via iterations. At each iteration, given a new remapped depth, the degree of visual discomfort is predicted again to determine if the predicted visual comfort score reaches a predefined target threshold and the depth differences between the connected planes are larger than the JNDD threshold simultaneously. If both conditions are satisfied, the depth optimization process is terminated. Otherwise, the process is repeated.

According to the experienced visual comfort level, it is natural to deem that for those visual comfortable planes, depth sensation can be further promoted by expanding the depth difference between planes (a higher value of γ_i should be selected), while for those visual uncomfortable planes, visual comfort is the primary factor in adjusting the depth range (a comparatively lower value of γ_i should be selected). Therefore, we further impose the following constraint for the selection of parameter γ_i

$$\gamma_i \succ \begin{cases} \text{larger than 1, if } \Phi(\mathbf{f}_i^o) \geq VC_{GT} \\ \text{lower than 1, otherwise} \end{cases} \quad (13)$$

In our implementation, an optimal value of γ_i is selected within a certain search range (with interval of 0.05). However, separately applying (12) for each plane may induce depth inversion (e.g., the adjusted depth range of one plane may be remapped to the depth range of another plane). To ensure that we remap our depth ranges exactly into the target space, we restrict the parameter γ_i to satisfy the following constraints:

$$\sum_{i=1}^L \gamma_i \leq \gamma_{global} \quad (14)$$

Here, γ_{global} denotes the global scale coefficient of depth range after global depth mapping optimization. Figure 2 shows a result after our two-stage depth mapping optimization. The left side shows results of original depth map. Our results are shown on the right. The histograms show that our approach nicely adjusts depth range for some discomfort regions (e.g., the girl), leading to an enhanced depth experience.

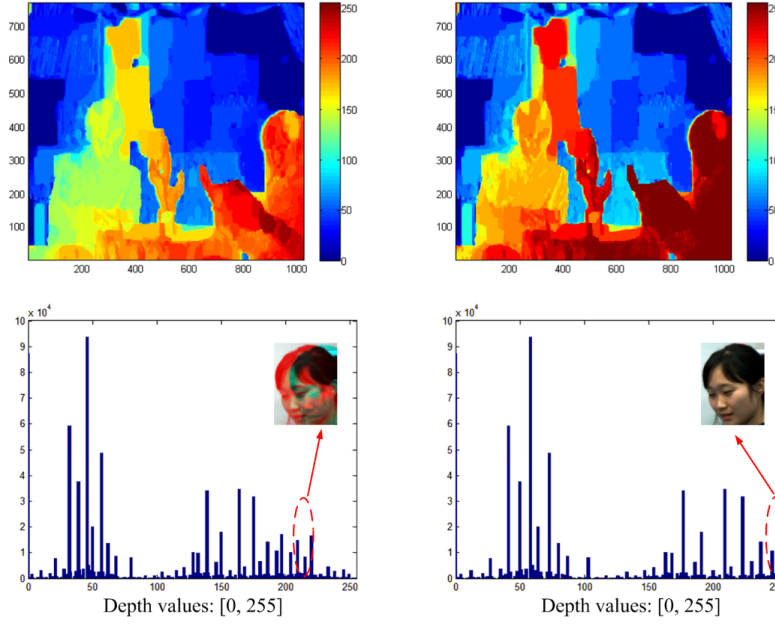


Fig. 2. Results of original (left) and our depth mapping. The histograms show how our approach nicely adjusts depth range for some discomfort regions.

3.4. Stereopair generation through 3D warping

Our final goal is to generate a stereopair from the remapped depth map. Considering the relationship between the camera coordinate and world coordinate, reference image is firstly mapped into world space and then projected to target image plane through 3D warping. The process can be separated into two steps: projection of the reference image into the 3D world coordinates, and followed by projection of the 3D scene into the target image plane. A pixel (u, v) in the reference image is projected into the 3D world coordinate (x, y, z) [6]:

$$(X, Y, Z)^T = \mathbf{R}_1 \mathbf{A}_1^{-1} (u, v, 1)^T d_{u,v} + \mathbf{T}_1 \quad (15)$$

where $d_{u,v}$ is the depth value calculated from the pixel (u, v) in the depth map, \mathbf{A}_1 (3×3) and \mathbf{R}_1 (3×3) are the intrinsic and rotation matrices of the reference camera, and \mathbf{T}_1 (3×1) is the translation vector of the reference camera. In the next step, the world coordinates are projected into the target camera plane via

$$(u', v', w')^T = \mathbf{A}_2 \mathbf{R}_2^{-1} (X, Y, Z)^T - \mathbf{A}_2 \mathbf{R}_2^{-1} \mathbf{T}_2 \quad (16)$$

where (u', v', w') is the homogeneous coordinates of the target image plane, and \mathbf{A}_2 , \mathbf{R}_2 and \mathbf{T}_2 are the intrinsic matrix, rotation matrix and translation vector of the target camera, respectively. The corresponding pixel location in the synthesized image of the target camera is $(x', y') = (u'/w', v'/w')$. The matrices \mathbf{A}_1 , \mathbf{R}_1 and \mathbf{T}_1 , as well as \mathbf{A}_2 , \mathbf{R}_2 , \mathbf{T}_2 , are known in advance for specific cameras. Finally, another hole filling operation is conducted to fill the remaining holes in the warped image.

4. Experimental results

In this section, we evaluate the proposed approach on some uncomfortable stereoscopic images ($\text{MOS} < 3$) and comfortable stereoscopic images ($\text{MOS} > 4$), to demonstrate that our depth mapping approach can effectively optimize the visual comfort and depth sensation to enhance the 3D viewing experience. To comprehensively evaluate the performance of the

proposed approach, we select three existing schemes for comparison: 1) Lei's scheme [13] that adjusts the disparity range by controlling the ZDP; 2) Jung's scheme [18] that constructs a saliency-adaptive nonlinear disparity mapping function; 3) our previous scheme [14] that adjusts the ZDP for projection.

4.1 Results of uncomfortable stereoscopic images

We analyze the effectiveness of the proposed approach on uncomfortable stereoscopic images. Figure 3 and Fig. 4 show comparison results of Lei's scheme, Jung's scheme, Shao's scheme and our approach on the No. 106 and No. 112 test images in the IVY LAB Stereoscopic 3D image database (even the database only provide disparity maps, we first convert them to depth maps based on camera parameters for our purpose). Viewing these images in 3D using red-green glasses demonstrates that the perceived visual comfort in (c) is much stronger than the results in (b), (d) and (e), but the perceived depth sensation is still low (the subjective assessment results in the following subsection 4.3 also demonstrate the conclusion). Compared with the results in (b), (c) and (d), the experienced visual comfort and depth sensation of our results in (e) is the highest. Since Lei's scheme and Shao's scheme both adjust the disparity ranges of the crossed and uncrossed disparities but maintain the entire disparity range, a negative disparity may be inverted to a positive value in some cases, while Jung's scheme scales the entire disparity range so that the disparities of all pixels are adjusted regardless of depth factors. In contrary, our approach will greatly diminish the depth sensation for those extremely uncomfortable regions (partly marked by blue ellipses), but still maintain enough depth sensation for other important regions. Another result in Fig. 5 for the 'Newspaper' 3D test sequence also demonstrates the conclusion (the estimated depth maps are provided in the sequence). Generally speaking, our approach can maintain a good tradeoff between enhancing visual comfort and reducing depth sensation.

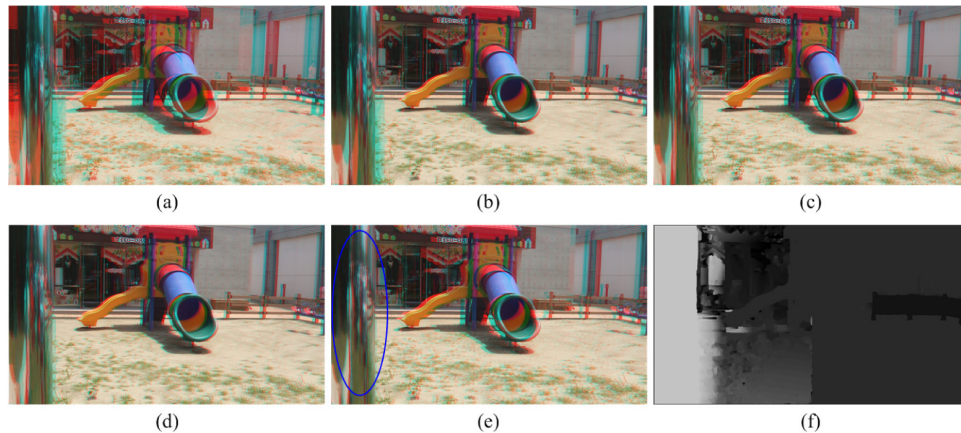


Fig. 3. Results of the No. 106 test image in the IVY LAB Stereoscopic 3D image database: (a) the anaglyph image of input stereo pair, (b) the anaglyph image of Lei's scheme [13], (c) the anaglyph image of Jung's scheme [18], (d) the anaglyph image of Shao's scheme [14], (e) the anaglyph image of our approach, and (f) the adjusted depth map of our approach. The anaglyph images should be viewed with a red-green glass.

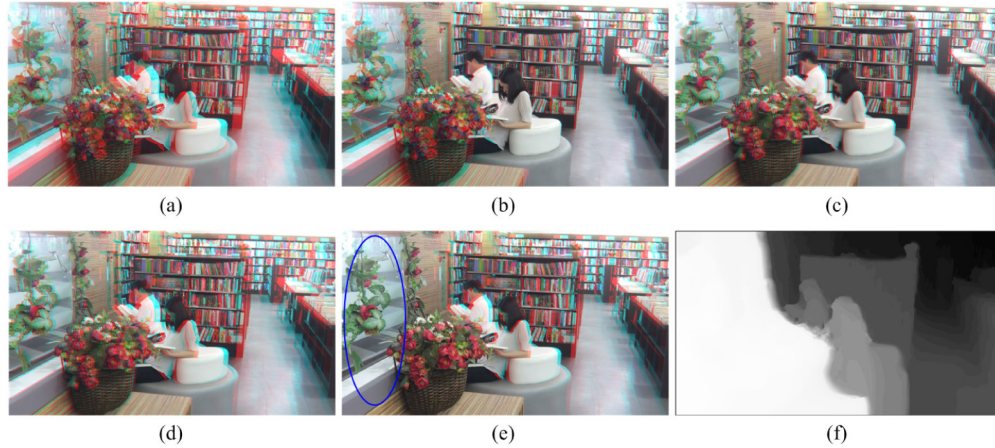


Fig. 4. Results of the No. 112 test image in the IVY LAB Stereoscopic 3D image database: (a) the anaglyph image of input stereo pair, (b) the anaglyph image of Lei's scheme [13], (c) the anaglyph image of Jung's scheme [18], (d) the anaglyph image of Shao's scheme [14], (e) the anaglyph image of our approach, and (f) the adjusted depth map of our approach. The anaglyph images should be viewed with a red-green glass.

4.2 Results of comfortable stereoscopic images

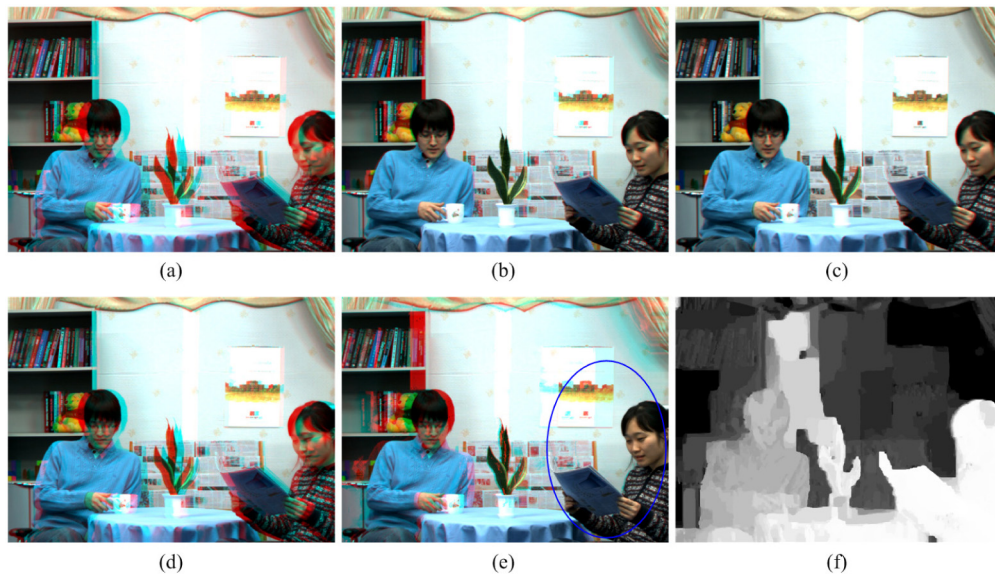


Fig. 5. Results of the 'Newspaper' 3D test sequence: (a) the anaglyph image of input stereo pair, (b) the anaglyph image of Lei's scheme [13], (c) the anaglyph image of Jung's scheme [18], (d) the anaglyph image of Shao's scheme [14], (e) the anaglyph image of our approach, and (f) the adjusted depth map of our approach. The anaglyph images should be viewed with a red-green glass.

Even though our approach is designed to enhance visual comfort for those uncomfortable stereoscopic images, it is worthwhile to examine if our approach will be still effective when performed on those comfortable stereoscopic images. The comparison results for the No. 115 test image in the IVY LAB Stereoscopic 3D image database in Fig. 6 show that, that the perceived visual comfort in (a), (b), (d), (d) and (e) all belong to a comfortable scale, but the perceived depth sensation in (e) is obviously higher than all other results (the subjective

assessment results in the following subsection 4.3 also demonstrate the conclusion). Since depth sensation is not carefully taken into account in Lei's scheme, Jung's scheme and Shao's scheme, their performances on those uncomfortable stereoscopic images are limited. In fact, in the practical applications, the degree of visual discomfort for a stereoscopic image is usually unknown. From this aspect, our approach may provide a simultaneous visual comfort and depth sensation optimization solution for stereoscopic images.

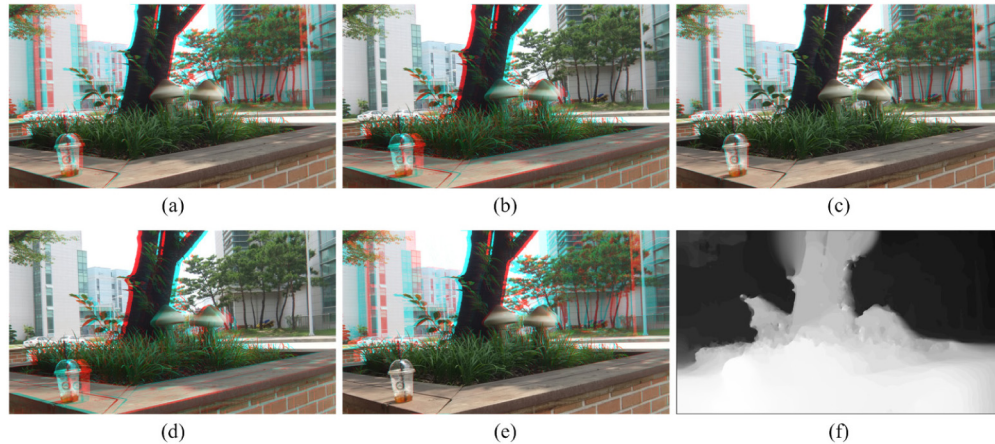


Fig. 6. Results of the No. 115 test image in the IVY LAB Stereoscopic 3D image database: (a) the anaglyph image of input stereo pair, (b) the anaglyph image of Lei's scheme [13], (c) the anaglyph image of Jung's scheme [18], (d) the anaglyph image of Shao's scheme [14], (e) the anaglyph image of our approach, and (f) the adjusted depth map of our approach. The anaglyph images should be viewed with a red-green glass.

4.3. Subjective assessment results

We validated our method using subjective testing. The subjective testing was conducted in the laboratory designed for subjective quality tests according to the recommendations ITU-R BT.500-11 [31] and ITU-R 1438 [32]. All stereoscopic images were randomly displayed on Samsung UA65F9000 65 inch Ultra HD 3D-LED TV. 3D Shutter Glasses were used in the subjective experiments. A single-stimulus Absolute Category Rating (ACR) test methodology described in ITU-T P.910 [33] and ITU-T P.911 [34] was used in the experiment. Eighteen graduate students participated in this subjective evaluation. All participants in the experiment met the minimum stereo acuity less than 60 seconds of arc (sec-arc), and passed a color version test. During the testing, the participants were asked to rate the stereoscopic images based on their experienced visual comfort on a five-level (ACR-5) scale: 5 = very comfortable, 4 = comfortable, 3 = mildly comfortable, 2 = uncomfortable and 1 = extremely uncomfortable, and the experienced depth sensation on a ACR-5 scale: 5 = very high, 4 = high, 3 = medium, 2 = low and 1 = very low. After detecting and discarding outliers of all opinion scores, the final MOS for each stereoscopic image was calculated as the mean of the remaining opinion scores.

The subjective assessment results on six uncomfortable stereoscopic images (three pairs are selected from the IVY LAB Stereoscopic 3D image database and three pairs are selected from 3D test sequence), and six uncomfortable stereoscopic images (all selected from the IVY LAB Stereoscopic 3D image database), are presented in Table 1 and 2. For these uncomfortable stereoscopic images, our approach can achieve a significant visual comfort improvement reaching to a comfortable scale. For those comfortable stereoscopic images, our approach may appropriately maintain a comfortable scale but promote the depth sensation. On the other hand, depth sensation for these uncomfortable stereoscopic images is appropriately reduced but still maintained in a high scale. In fact, for these uncomfortable stereoscopic images, the perceived depth is mainly from those perceptually uncomfortable

objects. If the perceived discomfort is reduced for those objects, the perceived depth is also reduced. Overall, our approach can achieve a better tradeoff between the visual comfort and depth sensation for arbitrary stereoscopic images.

Table 1. Quantitative Subjective Assessment Results of Different Stereoscopic Images













Method							
VC	Org	3.07	3.87	3.20	3.40	3.33	2.40
	Lei [13]	3.53	3.87	3.60	3.60	3.53	3.53
	Jung [18]	4.07	4.40	4.20	4.13	4.27	4.27
	Shao [14]	3.60	3.87	3.80	3.67	3.87	4.33
	Our	4.13	4.27	4.13	4.07	4.13	4.00
DS	Org	4.13	3.87	3.60	4.27	4.13	4.47
	Lei [13]	4.33	3.60	3.53	4.33	4.20	4.13
	Jung [18]	3.27	3.20	3.47	3.33	3.27	2.93
	Shao [14]	4.40	3.87	4.00	4.27	4.13	3.67
	Our	4.00	4.00	4.07	4.07	4.00	3.87

Table 2. Quantitative Subjective Assessment Results of Different Stereoscopic Images

Method							
VC	Org	4.20	4.40	4.27	4.53	4.07	4.07
	Lei [13]	4.07	4.33	4.20	4.47	4.13	4.20
	Jung [18]	4.20	4.53	4.40	4.67	4.33	4.13
	Shao [14]	4.13	4.27	4.07	4.20	4.07	4.13
	Our	4.13	4.27	4.13	4.13	4.00	4.00
DS	Org	3.27	3.13	3.13	2.67	3.27	3.07
	Lei [13]	3.73	3.40	3.40	2.87	3.53	3.33
	Jung [18]	3.40	3.07	3.00	3.13	3.13	3.27
	Shao [14]	3.87	3.67	3.73	3.40	3.60	3.40
	Our	3.93	3.93	4.00	4.13	4.07	4.07

4.4. The impact of two-stage optimization

In our two-stage depth optimization solution, at the first stage, the depth maps are processed using the global depth mapping optimization approach; at the second stage, the depth maps are further processed using the local depth mapping optimization approach. For clarity, the two approaches are denoted by Scheme-A and Scheme-B, respectively, and their independent performances are analyzed in the subsection. The comparison results on the No. 106 and No. 115 test images in Fig. 7 show that, for the uncomfortable stereoscopic image (the first row in the figure), the result only with Scheme-A or Scheme-B cannot diminish the visual discomfort for those extremely uncomfortable regions, while our two-stage approach can solve this problem. For the comfortable stereoscopic image (the second row in the figure), the result only by Scheme-B will largely reduce the perceived depth sensation. Overall, fusing the global and local depth mapping optimization operation will produce satisfactory results. Even though our approach is primarily designed to tackle with uncomfortable stereoscopic images, it is still effective for those comfortable stereoscopic images, because the visual comfort and depth sensation are simultaneously enhanced via two-stage global and local optimization.

From this aspect, our approach may provide a ‘completely blind’ solution for stereoscopic images.



Fig. 7. Results of the two-stage optimization: (a) the anaglyph image of Scheme-A, (b) the anaglyph image of Scheme-B, and (c) the anaglyph image of our approach.

4.5. Discussions

- 1) *The visual comfort prediction function*: Up to now, there is no model that can predict the degree of visual discomfort for stereoscopic images in all cases, since many causes (e.g., excessive disparity gradient, small stimulus width, high spatial frequency, etc.) are related to visual discomfort. Even though the visual comfort prediction function is effective in this paper, whether or not this function is still suitable for other databases is still an open problem worth discussing. This general problem is also existed in other visual comfort enhancement methods.
- 2) *The accuracy of depth map*: The state-of-the-art depth sensation enhancement methods are dependent on the ground-truth depth maps to separate the depth layers. However, accurate disparity/depth estimation is still an open problem in computer vision. The proposed method may fail to characterize the depth differences when the extracted depth layers are incorrect, leading to an unnatural depth sensation. Even though disparity estimation is not the focus of this paper, it is still worth exploring how to solve the problem through some other ways.

5. Conclusions

We have presented a general framework for depth mapping to optimize visual comfort for S3D display based on color-plus-depth signals. The most important technical innovation of our framework is that several factors that affect visual comfort and depth sensation are integrated, and an optimization solution is provided to remap the depth range of color-plus-depth signal for S3D display. Our results demonstrate the effectiveness and usability of our approach. Our work is a starting point to tackle this challenging problem, and leaves a lot of room for improvement.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) (grant 61271021, 61471212, 61271270 and U1301257), the National High-tech R&D Program of China (grant 2015AA015901). It was also sponsored by K.C.Wong Magna Fund in Ningbo University.