

S3D-R2R: An Automatic Stereoscopic 3D Image Recomposition to Retargeting Method with Depth Modification

Md. Baharul Islam^{1,2}^a, Chee Onn Wong² and Md. Kabirul Islam³

¹American University of Malta, Bormla 1013, Malta

²Multimedia University, Cyberjaya 63100, Malaysia

³Daffodil Int. University, Dhaka 1207, Bangladesh

Keywords: Disparity, Image Aesthetics, Image Recomposition, Image Retargeting, Image Warping, Stereoscopic Imaging.

Abstract: Stereoscopic image adaptation to the target display devices while minimizing the distortion of significant features and stereoscopic properties is a challenging problem. Conventional methods either fail to preserve the image context or unable to improve the image aesthetics with improved depth perception in the retargeted images. In this paper, we present an automatic warping-based *stereoscopic 3D image recomposition to retargeting method*, shortly **S3D-R2R** that improves the stereo image composition in the retargeting results. Our **S3D-R2R** method resizes both the left and right stereo image pair using a global optimization algorithm that minimizes a set of aesthetic quality errors. These errors are formulated based on the selected photographic composition rules and modify the depth perception. To improve the depth perception of the stereo image pair, the disparity consistency has been modified within the comfort disparity range. Experimental results show that our automatic method changes the position of the salient object in the target image scale and improves the depth perception within the comfort depth range. Empirical user studies indicate that our retargeting results receive more attention than state-of-the-art methods.

1 INTRODUCTION

Due to the rapid growth of 3D display devices, stereoscopic 3D image and video contents are widely available in online. The 3D viewing experience can vary on different display devices with various aspect ratios and sizes. Conventional monocular image retargeting methods did not consider the stereoscopic properties when applied independently to the stereoscopic left and right image pair. These methods can create a discomfort 3D viewing experience and may result the eyestrain and headache. Stereoscopic image retargeting requires additional attention for avoiding these 3D fatigues. The naive scaling of the stereoscopic image distorts the objects shape and can be responsible for the unpleasant 3D viewing experience. The black box solution wastes the free space of the images. Stereo cropping (Niu et al., 2012b) utilizes the cropping operator on both the left and right stereo image based on some selected single and stereo photographic composition rules. This method works well

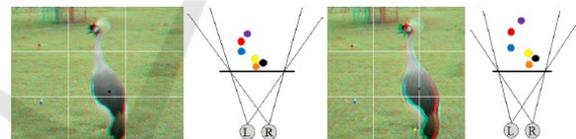


Figure 1: An example of our **S3D-R2R** result; (Left to right) original anaglyph (red-cyan) image with some feature points (white lines represent the rule of thirds composition), depth distribution of those feature points in 2D space (*L* and *R* represent the left and right eye view), retargeting result in our **S3D-R2R** method (image width reduced by 20%), and depth distribution of the corresponding feature points.

with the sufficient uninteresting background. However, image cropping suffers the information loss and may not be produced a good results when the salient objects are spread all over the image frame. Besides, the depth perception can be reduced if it is aggressively cropped-off.

Content-driven warping (Yoo et al., 2013) preserved the disparity between the left and right stereo image pair while object-coherence warping (Lin et al., 2014) additionally used a shape preservation constraint to protect the shape of the salient objects.

^a <https://orcid.org/0000-0002-9928-5776>

However, these methods did not consider the image aesthetics in the retargeting process. Aesthetics-driven warping (AWARP) (Islam et al., 2015) proposed to retarget the stereoscopic image pair based on the photo composition while preserving the disparity in their results. Although this method can change the spatial position of the foreground objects, but it is also desirable to modify depth perception within the comfort depth zone of poorly taken (particularly, with lower depth images) stereo photographs by amateurs.

To address this limitation, we propose an automatic, warping-based method that can enhance 3D viewing experience in the retargeted stereo images by modifying composition and depth. Our **S3D-R2R** method minimizes aesthetic quality errors that are formulated based on the photo composition using a global optimization algorithm. The depth modification within the comfort depth range is considered as an optimization problem in our method. Figure 1 shows an example of our retargeting result. The salient object (bird) changes the optimal position according to the rule of third composition. The depth perception of the selected feature points (marked as color dots) is also enhanced than the original counterpart. The best view can be used as an anaglyph (red-cyan) glass to perceive the depth information in the color version.

2 RELATED WORKS

A numerous single image retargeting methods have been proposed from the last decade. The independent application of these methods to the left and right stereo image pair can easily destroy the stereoscopic properties that cause eyestrain and headache. The subjective (Islam et al., 2017) and/or objective evaluation (Ma et al., 2012) of these methods had been conducted. In this section, we only provide a brief summary of the non-aesthetic and aesthetic-based stereoscopic image retargeting, recomposition and depth enhancement methods.

2.1 Stereoscopic Image Retargeting

Stereo cropping (Zhang et al., 2013) utilized the cropping operator to resize the stereo image pair. This method ensures to avoid stereoscopic violation. However, it was responsible to loss the information of the resized image. Seam carving on stereo image (Basha et al., 2011) extended the seam carving on single image retargeting. The seam is a interconnected pixels from top to bottom or left to right in the image. They remove a pair of seams from both

the left and right stereo image pair for getting the resized stereo image pair. Recently, seam carving also applied on the stereoscopic video retargeting (Guthier et al., 2013) that used an additional constraint to maintain temporal consistency between two consecutive video frames. The noticeable features and geometric distortions are visible in seam carving based methods due to its discrete nature. A stereo image with large disparity, shift-map (Qi and Ho, 2013) integrated the retargeting results with depth adjustment simultaneously for reducing the large disparity of the original images. This method may suffer geometric and semantic distortions in the results.

Continuous methods generally optimize a set of triangular/quad meshes, subject to a set of constraints. Content-aware methods (Li et al., 2015) used the monocular warping-based image resizing to stereoscopic domain that aims to preserve stereoscopic properties using two stereoscopic constraints; vertical alignment for avoiding vertical artifact, and disparity consistency between the left and right stereo image pair. Niu et al. (Niu et al., 2012a) proposed an enable warping to resize a stereo image which has a clear objective to preserve prominent objects with its 3D structure. Recently, a warping-based stereo video retargeting (Islam et al., 2019) has been proposed that ensured the temporal consistency between two consecutive video frames in the retargeted video.

Scene warping (Lee et al., 2012) decomposed the given image into several layers according to the depth orders and each layer is warped according to its own mesh deformation. The warped layers were then composited together according to depth order to get the retargeted images. This method ensures object protection, but it may not able to ensure the semantic connectedness (e.g. shadow) between the foreground object and its background environment. Then, semantic preserving warping (Tan et al., 2015) has been proposed to overcome this limitation. This method ensured to protect objects, correct depth order, and semantic connectedness between foreground objects and its immediate background.

2.2 Recomposition to Retargeting

The unpleasant stereo image (due to poor composition with lower depth perception) can be aesthetically pleasing by using the aesthetics-driven stereo image recomposition to retargeting. An automatic stereo cropping (Niu et al., 2012b) utilized the cropping operator to recombine both the left and right stereo images based on the some single and stereoscopic photographic composition rules. This approach may suffer content loss and not useful when a single or multi-

ple objects are spread out a significant portion in the image frame. Then, AWARP (Islam et al., 2015) preserved the global image context. Both of the above methods only allowed for changing the optimal position of the salient objects and preserving the disparity in their results. Due to preserving disparity, the 3D viewing experience of the retargeting results is similar to the original stereo images in these methods. Recently, a hybrid stereoscopic image recombination (Islam et al., 2018) method has been proposed that not only modify the spatial composition but also depth remapping.

2.3 Depth Modification

The cause behind the visual discomfort and 3D fatigue (e.g. excessive display screen, accommodation and convergence/divergence mismatch, out of comfort depth zone) of stereoscopic images are details discussed in (Lambooij et al., 2007). Recently, some promising automatic and interactive depth enhancement methods have been proposed. A warping-based method (Du et al., 2013) used to change the perspective of the stereo images through some advanced camera effects such as dolly zoom and wide angle effects. The shift-map stereo image editing (Yan et al., 2013) adjusted the depth (especially for images with large disparity) and preserves the 3D scene structure in their results. All of these methods only considered to enhancing the depth perception without retargeting and recombination. 3D Copy&Paste (Lo et al., 2010) is an end-to-end billboard system that segments the objects from a source stereo image pair and paste back to the target stereo image pair while preserves the stereoscopic properties. StereoPasting (Tong et al., 2013) slightly improved the stereo 3D Copy&Paste that did not require the input stereo image pair. It segments the foreground objects from a 2D image and then pastes the segmented objects onto the 3D background scene.

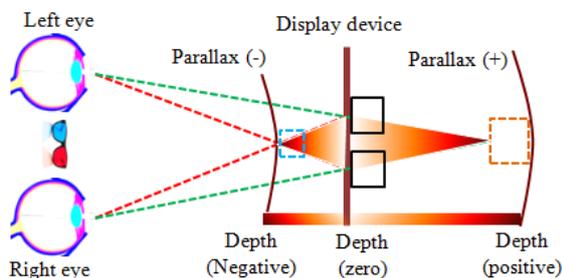


Figure 2: An overview of the comfort depth zone in the binocular vision system. The object appears in front of the display device (blue dot rectangle) for negative parallax (red dot lines) and behind the screen (yellow dot rectangle) for positive parallax (green dot lines). The brown arcs represent the comfort depth perception range in human vision system.

3 COMFORT DEPTH ZONE

Due to horizontally separated of the human eyes by 65mm (for adult), they perceive two slightly different images (in 2D space) of the same scene from the left and right view. These two 2D images are fused into the human brain to perceive the depth information. The difference between the left and right image is called *binocular disparity*. An object can appear in front and/or behind the screen in the 3D virtual world depending on the nature of disparity (negative/positive parallax). The human vision system comfortably perceives a limited amount of depth information, namely *comfort depth zone/range*. The disparity more than human interpupillary distance (65 mm) defuses the left and right view in human brain and may result discomfort 3D viewing experience and 3D fatigue. The depth perception of the binocular image also depends on the distance between the viewer and the location of the display device (Mendiburu, 2012). Figure 2 shows the stereoscopic comfort depth range of the human binocular vision system. The *square* is a binocular object, perceived by both left and right eyes with negative parallax (red dots lines) on the display device. The object (blue square) can appear in front of the display device in the 3D virtual world. After shifting the left and right views, the object (orange square) can appear behind the display device due to positive parallax on the display device (green dots lines). The state-of-the-art automatic stereo retargeting methods relocate the salient objects and preserve the disparity for getting similar 3D viewing experience as an original stereo image. The object relocation may be not enough for the aesthetics-driven retargeting, especially for the stereo image with lower depth perception. In this **S3D-R2R** method, we modify the depth perception within the comfort depth zone with changing the object position.

4 S3D-R2R METHOD

The aim of our **S3D-R2R** method is to retarget both left and right stereo images within the target image scale for enhancing image aesthetics and modify depth perception of the retargeted images. Figure 3 shows an overview of the **S3D-R2R** method. Given a stereo image pair, we first compute the Sum of Absolute Difference (SAD) between the left and right stereo image. Then, the triangular meshes are formulated over both the left and right image pair based on the SAD. In the second stage, we minimize a set of errors includes warping, aesthetic quality, and stereo errors, subject to a set of constraints.

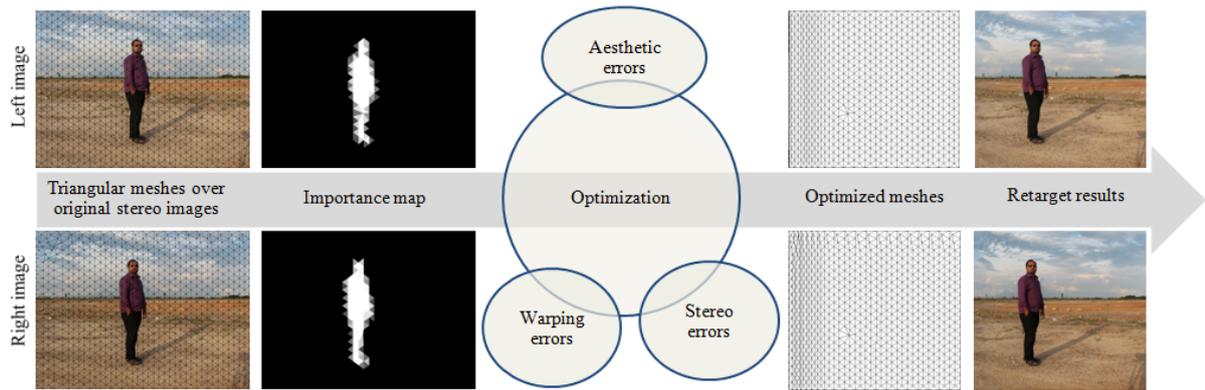


Figure 3: An overview of our proposed S3D-R2R method. It has mainly two steps: (1) significance mesh computation, (2) warping based minimization of a set of errors, subject to a set of constraints.

4.1 Significance Mesh Computation

The SAD between the left, I^L and right, I^R stereo image pair is calculated using equation 1. Delauney triangular meshes, M^L is constructed to represent the left stereo image I^L . We employ the simplified visual saliency (Harel et al., 2006) on I^L . A pixel with high saliency value is considered as a significant image pixel. The corresponding triangular meshes, M^R and saliency to the right image I^R is then automatically propagated based on the SAD information. A set of pixels in a triangle with high saliency is considered as a significant triangle. Figure 2 shows the significant triangular meshes, $S_L = \{s_1, s_2, \dots, s_n\}$ from $M^L \cup M^R$, where n is the total number of significant triangular meshes. In order to avoid the distortion of salient objects, the significant triangles keep as rigid as possible during the optimization process.

$$d = \sum_{(x,y) \in W} \|I^L(x,y) - I^R(x+i,y+j)\| \quad (1)$$

where (x,y) is representing the pixel location in the left image I^L while $(x+i,y+j)$ is the corresponding pixels in I^R . W is the window size in $I^L \cup I^R$.

4.2 Non-homogeneous Warping

The left and right meshes, $M^L \cup M^R$ constrain a set of vertex, $\mathbf{V} = \{v_1, v_2, \dots, v_m\}$. In the warping process, the source triangular meshes, $M^L \cup M^R$ are mapped to the target meshes $\bar{M}^L \cup \bar{M}^R$ respectively. Let, the set of triangle meshes $\mathbf{T} = \{t_1, t_2, \dots, t_t\}$, set of significant triangles $\mathbf{S} = \{s_1, s_2, \dots, s_n\}$, and set of objects $\mathbf{O} = \{o_1, o_2, \dots, o_o\}$. Ideally, the significant triangles in $M^L \cup M^R$ could be homogeneously and other triangles could be non-homogeneously scaled along with x - and/or y -directions without rotation.

4.2.1 Warping Errors

The warping errors consist the scale transformation and smoothness error. For each triangle $t \in T$, we perform non-uniform scaling s_x and s_y in x - and y -direction, respectively. To avoid the discontinuity between two neighbouring triangles t and s , we constrain mesh transformation smoothly to the target mesh, $M^L \cup M^R$. The scale transformation, E_w and smoothness error, E_s are defined as,

$$E_w = \sum_{t \in T} A_t \|J_t - G_t\|_F^2 \quad (2)$$

$$E_s = \sum_{t,s \in T} A_{st} \|G_t - G_s\|_F^2 \quad (3)$$

where A_t is the area of triangle t , $\|\cdot\|_F^2$ is the Frobenius norm, J_t is a 2×2 Jacobian matrix that maps a triangle to its corresponding triangle in the output mesh $\bar{M}^L \cup \bar{M}^R$, $A_{st} = (A_s + A_t)/2$, and s, t are adjacent triangles.

4.2.2 Recomposition Errors

Employing photographic composition rules to stereo images can enhance image aesthetics in the retargeting results. We apply two photo composition rules in our S3D-R2R method. Besides, we also consider the depth modification within the comfort depth range in the optimization process.

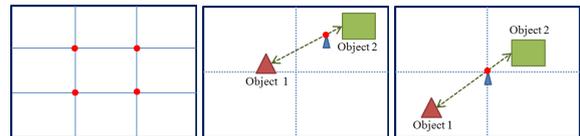


Figure 4: Rule of thirds and visual balance composition; (Left to right) the rule of thirds composition, four red dots are representing the power points, unbalanced composition of two objects (triangle and rectangle), and visually balanced composition.

Rule of Thirds Error: In this rule, the image frame is divided by two vertical and two horizontal lines that create four intersection points, namely *power points*. Photographers are encouraged to place the center of mass of most salient objects on these points. Figure 4 shows the rule of thirds composition. In our **S3D-R2R** method, we minimize the distance between the power point and the center of salient objects. The rule of thirds error is defined as,

$$E_p = \sum_{o \in O, s \in S} A_s \|D_p - O_o\| \quad (4)$$

where D_p is the power points, O_o centroid of objects, A_s is the area of important triangle s .

Visual Balance Error: The center of the visual mass of all salient objects should be placed on the image center that creates harmony between objects. Figure 4 shows an unbalanced and a balanced composition for two salient objects. The visual balance error, E_{vb} is defined as,

$$E_{vb} = \sum_{o \in O} A_s \|C(I) - C(O)\| \quad (5)$$

where $C(I)$ is the image center, $C(O_o)$ is weighted centroid of salient objects O , and A_s is the area of important triangle s . The visual balance error, $E_{vb} = 0$ for the image with only one object.

4.2.3 Depth Modification

The depth perception depends on the disparity and the distance between the position of the viewers and the display device (discussed in Section 2.3). Firstly, we change the disparity from pixel domain to the physical domain by dividing the pixel density of our display device. The boundary vertices, \mathbf{B} in input meshes, $M^L \cup M^R$ are constrained to the boundary vertices in the output meshes, $\bar{M}^L \cup \bar{M}^R$. The number of the optimizable vertices, $\mathbf{N} = \mathbf{V} - \mathbf{B}$. If the SAD (as disparity) of a particular vertex $v \in \mathbf{N}$ is d^v , then the average depth perception, E_z is defined as,

$$E_z = \frac{1}{N} \sum_{v \in \mathbf{N}} \frac{eD}{e - d^v} \quad (6)$$

$$E_n = s_z E_z \quad (7)$$

where e is the interpupillary distance between human eyes, D is the distance between the viewer and the display device, and s_z is the depth modification scale. We set $e = 6.5$ and $D = 100$ in our experiment. The comfort depth range, R is set to 78 – 140cm (Lambooi et al., 2007) in the physical domain for the above setting.

4.2.4 Stereoscopic Quality Error

It is required to minimize the changes between the left and right warped meshes $\bar{M}^L \cup \bar{M}^R$ for avoiding 3D fatigue. Let, (v^L, v^R) and (\bar{v}^L, \bar{v}^R) denote the set of the corresponding vertex of the input meshes $M^L \cup M^R$ and output meshes, $\bar{M}^L \cup \bar{M}^R$ respectively. For each $v \in \mathbf{N}$, we minimize the changes of vertical alignment between the left and right meshes, $\bar{M}^L \cup \bar{M}^R$. The vertical alignment error is defined as,

$$E_v(v^L, v^R) = \|\bar{v}^R(y) - \bar{v}^L(y)\|^2 \quad (8)$$

where (y) refers the y coordinate values in $\bar{M}^L \cup \bar{M}^R$.

4.3 Error Minimization

The total error is formulated as the combination of warping, aesthetic, and stereoscopic quality errors.

$$E_T = w_w E_w + w_s E_s + w_p E_p + w_{vb} E_{vb} + w_n E_n + w_v E_v \quad (9)$$

where $w_w, w_s, w_p, w_{vb}, w_n$, and w_v are the corresponding weights of warping, aesthetic and stereoscopic quality errors.

The boundary vertices in $M^L \cup M^R$ keep as boundary vertices in $\bar{M}^L \cup \bar{M}^R$. For each boundary vertex $v \in \mathbf{B}$ of the input mesh $M^L \cup M^R$, we apply the boundary position constraint to the left, right, top and bottom border vertices, respectively. The total error function in Equation 9 is a convex quadratic function. We utilize the *cvx* optimization (Grant et al., 2008) to find the solution to the quadratic function. The warping error weights w_w and w_w are set to 1 and 0.5, the aesthetic error weights w_p, w_{vb} , and w_n are set to 0.5, 0.5 and 0.1, and the vertical alignment weight is set to 1 in our experiment respectively.

5 EXPERIMENTAL RESULTS

Our **S3D-R2R** method is tested on an Intel i7 CPU, 3.40GHz with 12GB memory. The computation time is about 2-6 seconds to obtain the recomposition to retargeting results. The computation time depends on both left and right triangular meshes and image resolution. For a stereoscopic image size 1024×681 , the computation time is 5.32s including significance mesh computation prior optimization.

5.1 Retargeting Results

Figure 5 shows the different recomposition to retargeting results of two stereo images with single and

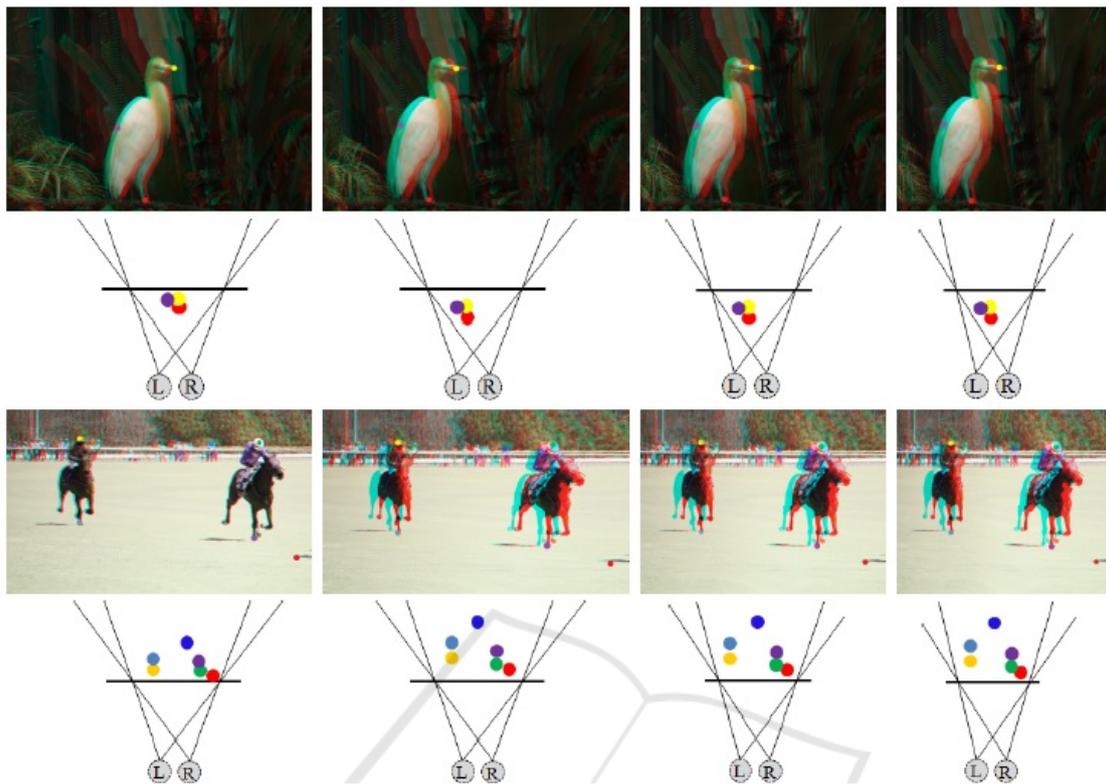


Figure 5: Different recomposition to retargeting results of our **S3D-R2R** method to a single and multiple salient objects; (Row 1, 3) anaglyph (red-cyan) image with some feature points, (Row 2, 4) depth distribution of selected feature points. (Left to right) original stereo image, recomposition result, recomposition to retargeting results by reducing image width at 20% and 30%, respectively.

multiple salient objects. The original stereoscopic images are recomposed and retarget by reducing the image width at 20%, and 30% respectively. Our method relocates the salient objects to an optimal position according to the composition rules: rule of thirds and visual balance. The single salient object (bird) is captured without following the rule of thirds composition in the original stereo image. In our recomposition and retargeting results, the salient object is relocated to the left power points. The depth perception range is also modified within the comfort depth range in our results. The bird in our results is closer to the viewer position than the original stereo image. Please refer to the depth distribution of the selected feature points (color dots). The user can perceive the depth by wearing an anaglyph (red-cyan) glass in the color version. In the second example, the salient objects (two horses) are visually unbalanced in the original stereo image. Besides, the depth perception of selected feature points is also poor (refer the selected feature points). Our results are visually balanced compared to the original stereo image. Besides, the depth perception range is also improved within the comfort depth range.

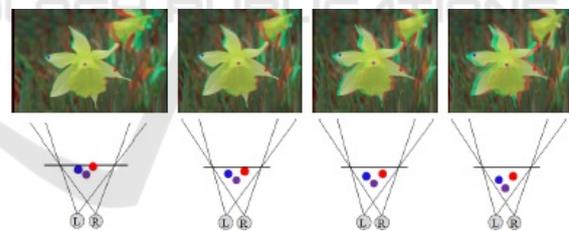


Figure 6: Retargeting results with different depth scale s_z ; (Left to right) original stereo image, the image width is reduced by 20% with $s_z = 0.01$, $s_z = 0.03$, and $s_z = 0.05$ respectively. (Top) anaglyph (red-cyan) image with three selected feature points, and (bottom) depth distribution of those feature points.

Figure 6 shows retargeting results with different depth scale s_z . The original anaglyph (red-cyan) image width is reduced by 20%. The salient object (flower) is appeared in front of the display device of the original anaglyph (red-cyan) image. Our retargeting result by setting $s_z = 0.01$ slightly improves the perceive depth range compared than original stereo image. The flower is appeared closer to the viewer position at $s_z = 0.03$ and $s_z = 0.05$, respectively. Best view may perceive by wearing an anaglyph (red-cyan) glass in the color version.

5.2 Comparison

We compare our recomposition to retargeting results with the state-of-the-art non-aesthetic based retargeting methods. Figure 7 shows the comparison of our method with non-aesthetic based retargeting methods. The stereo image width is reduced by 20%. Linear Scaling (LS) destroy the shape of the salient objects and reduce the depth perception in the resized image. The selected feature points are slightly moved towards the display device. We recommend readers either wearing anaglyph (red-cyan) glass or carefully follow the depth distribution of the selected feature points. The noticeable objects (two men) distortion are founded by geometrically consistent seam carving (Basha et al., 2011). This method may not be able to protect the salient objects due to its discrete nature. The content-driven stereo warping (Yoo et al., 2013) produces comparative better results than seam carving. For protecting foreground objects and scene consistency, semantic-preserving stereo warping (Tan et al., 2015) has been proposed to ensure the semantic connectedness between the foreground objects and its background layers. All these methods didn't consider the recomposition of the retargeted images. Besides, results of the state-of-the-art methods are visually unbalanced. In our method, the results not only change the optimal position of the salient objects (two men) and but also modify the depth perception within the comfort depth range. The depth distribution of the selected feature points is closer to the viewer than the state-of-art retargeting methods.

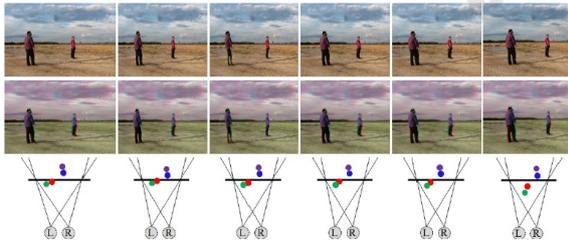


Figure 7: Compare our **S3D-R2R** result with the state-of-the-art non-aesthetics based retargeting results. (Top to bottom) stereoscopic left image, anaglyph (red-cyan) image with some feature points, and depth distribution of those feature points. (Left to right) original, linear scaling (LC), seam carving (SC) (Basha et al., 2011), traditional warping (WARP) (Yoo et al., 2013), semantic-preserving warping (TWARP) (Tan et al., 2015), and ours.

In Figure 8, we compare our result with state-of-the-art automatic aesthetic-driven retargeting results. Stereo cropping (Niu et al., 2012b) suffers the content loss. The salient object (climb man) is cropped off in the retargeting result. The AWARP (Islam et al., 2015) generates comparatively better results than the

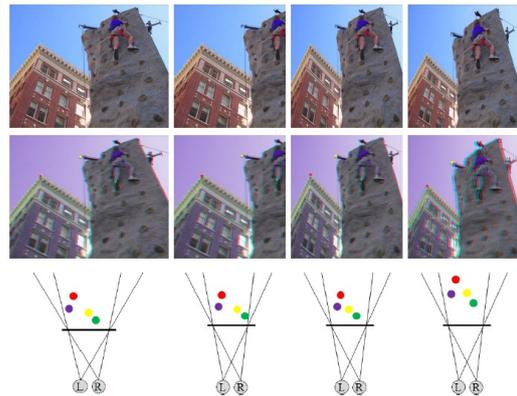


Figure 8: Compare our **S3D-R2R** method with the state-of-the-art aesthetics-driven retargeting methods. (Top to bottom) left stereo image, anaglyph (red-cyan) image with selected feature points, depth distribution of the selected feature points. The image width is reduced by 30%. (Left to right) original, results of stereo cropping (Niu et al., 2012b), AWARP (Islam et al., 2015) and ours.

CR and free from loss of information. Both of the above methods preserve the disparity consistency in their results. Our method modifies the depth perception within the comfort depth zone in the retargeted images. The selected feature points represents the more depth perception in our result.

5.3 Empirical User Study

Due to subjectivity of stereo image aesthetics, we conduct two empirical tests in our **S3D-R2R** method. We invite 30 independent subjects to compare our results with the AWARP (Islam et al., 2015). We observe that all the subjects had prior 3D viewing experience by watching 3D commercialized movies. We provide an NVIDIA GeForce 3D shutter glass to the subject. We test total 18 recomposition (without resizing) and 15 retargeting (through recomposition) results with single and multiple objects. We randomly display two sets of stereoscopic 3D images (AWARP, and our results) side by side and ask the subjects to pick the best retargeting results.

In our first study, we compare our recomposition results with AWARP. On average, subject's are preferred our results over the AWARP method at 17 out of 18 images (94.44%). For two recomposition results, 100% subjects prefer our results. In the second study, we compare our retargeting results with AWARP. On average, subject's prefer our results over the AWARP method 13 out of 15 images (86.66%). For two images (image no. 2 and 14), the results of AWARP method are preferred than our results.

6 CONCLUSION

In this paper, we present an automatic, recomposition to retargeting method for stereoscopic images using a global optimization algorithm, namely **S3D-R2R**. To maximize stereo image composition, we minimize a set of aesthetic quality errors formulated based on two photo composition rules during the warping process. Besides, our method can modify the depth perception in 3D space. It also minimizes the changes the vertical alignment between the left and right stereo image pair. Compared to stereo cropping and warping, our method can better preserve the global image context and able to modify depth perception for better 3D viewing experiences. The unavoidable feature distortions are found for the large scale warping, particularly stereoscopic images with complex/ geometric structures. Moreover, the aspect ratio of the salient objects can not be protected in our method. A shape preservation constraint and/or object segmentation can be used to solve this problem. In the future work, we would explore the stereoscopic video retargeting through recomposition.

REFERENCES

- Basha, T., Moses, Y., and Avidan, S. (2011). Geometrically consistent stereo seam carving. In *IEEE International Conference on Computer Vision (ICCV)*, pages 1816–1823.
- Du, S.-P., Hu, S.-M., and Martin, R. R. (2013). Changing perspective in stereoscopic images. *IEEE Transactions on Visualization and Computer Graphics*, 19(8):1288–1297.
- Grant, M., Boyd, S., and Ye, Y. (2008). Cvx: Matlab software for disciplined convex programming.
- Guthier, B., Kiess, J., Kopf, S., and Effelsberg, W. (2013). Seam carving for stereoscopic video. In *11th IEEE IVMS Workshop*, pages 1–4. IEEE.
- Harel, J., Koch, C., and Perona, P. (2006). A saliency implementation in matlab.
- Islam, M. B., Lai-Kuan, W., and Chee-Onn, W. (2017). A survey of aesthetics-driven image recomposition. *Multimedia Tools and Applications*, 76(7):9517–9542.
- Islam, M. B., Lai-Kuan, W., Chee-Onn, W., and Low, K.-L. (2015). Stereoscopic image warping for enhancing composition aesthetics. In *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, pages 645–649. IEEE.
- Islam, M. B., Wong, L.-K., Low, K.-L., and Wong, C.-O. (2018). Aesthetics-driven stereoscopic 3-d image recomposition with depth adaptation. *IEEE Transactions on Multimedia*, 20(11):2964–2979.
- Islam, M. B., Wong, L.-K., Low, K.-L., and Wong, C. O. (2019). Warping-based stereoscopic 3d video retargeting with depth remapping. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1655–1663. IEEE.
- Lambooi, M. T., IJsselstein, W. A., and Heynderickx, I. (2007). Visual discomfort in stereoscopic displays: a review. In *Electronic Imaging*, pages 649001–649001. International Society for Optics and Photonics.
- Lee, K. Y., Chung, C. D., and Chuang, Y. Y. (2012). Scene warping: Layer-based stereoscopic image resizing. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 49–56.
- Li, B., Duan, L.-Y., Lin, C.-W., Huang, T., and Gao, W. (2015). Depth-preserving warping for stereo image retargeting. *IEEE Transactions on Image Processing*, 24(9):2811–2826.
- Lin, S. S., Lin, C. H., Chang, S. H., and Lee, T. Y. (2014). Object-coherence warping for stereoscopic image retargeting. *IEEE Transactions on Circuits and Systems for Video Technology*, 24(5):759–768.
- Lo, W.-Y., van Baar, J., Knaus, C., Zwicker, M., and Gross, M. (2010). Stereoscopic 3D copy & paste.
- Ma, L., Lin, W., Deng, C., and Ngan, K. N. (2012). Image retargeting quality assessment: a study of subjective scores and objective metrics. *IEEE Journal of Selected Topics in Signal Processing*, 6(6):626–639.
- Mendiburu, B. (2012). *3D movie making: stereoscopic digital cinema from script to screen*. CRC Press.
- Niu, Y., Feng, W.-C., and Liu, F. (2012a). Enabling warping on stereoscopic images. *ACM Transactions on Graphics (TOG)*, 31(6):183.
- Niu, Y., Liu, F., Feng, W. C., and Jin, H. (2012b). Aesthetics-based stereoscopic photo cropping for heterogeneous displays. *IEEE Transactions on Multimedia*, 14(3):783–796.
- Qi, S. and Ho, J. (2013). Shift-map based stereo image retargeting with disparity adjustment. In *11th Asian Conference on Computer Vision (ACCV)*, pages 457–469.
- Tan, C.-H., Islam, M. B., Wong, L.-K., and Low, K.-L. (2015). Semantics-preserving warping for stereoscopic image retargeting. In *Image and Video Technology*, pages 257–268. Springer.
- Tong, R.-F., Zhang, Y., and Cheng, K.-L. (2013). StereoPasting: interactive composition in stereoscopic images. *IEEE Transactions on Visualization and Computer Graphics*, 19(8):1375–85.
- Yan, T., He, S., Lau, R. W., and Xu, Y. (2013). Consistent stereo image editing. In *Proceedings of the 21st ACM international conference on Multimedia*, pages 677–680. ACM.
- Yoo, J. W., Yea, S., and Park, I. K. (2013). Content-Driven Retargeting of Stereoscopic Images. *IEEE Signal Processing Letters*, 20(5):519–522.
- Zhang, F., Niu, Y., and Liu, F. (2013). Making stereo photo cropping easy. In *IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE.