Advanced Stereo Seam Carving by Considering Occlusions on Both Sides

YONGYI GONG, SHANGRU LI, and KANOKSAK WATTANACHOTE, Guangdong University of Foreign Studies, China
XIAONAN LUO, Guilin University of Electronic Technology, China

Stereo image retargeting plays a significant role in the field of image processing, which aims at making major objects as prominent as possible when the resolution of an image is changed, including maintaining disparity and depth information at the same time. Some seam carving methods are proposed to preserve the geometric consistency of the images. However, the regions of occlusion on both sides are not considered properly. In this article, we propose a solution to solve this problem. A new strategy of seams finding is designed by considering occluded and occluding regions on both of the input images, and leaving geometric consistency in both images intact. We also introduced the method of line segment detection and superpixel segmentation to further improve the quality of the images. Imaging effects are optimized in the process and visual comfort, which is also influenced by other factors, can be boosted as well.

CCS Concepts: • **Computing methodologies → Computer vision**; *Image processing*;

Additional Key Words and Phrases: Stereo image retargeting, occluded regions, occluding regions, avoiding strategy, line segment detection

ACM Reference format:
Yongyi Gong, Shangru Li, Kanoksak Wattanachote, and Xiaonan Luo. 2019. Advanced Stereo Seam Carving by Considering Occlusions on Both Sides. *ACM Trans. Multimedia Comput. Commun. Appl.* 15, 3, Article 69 (August 2019), 21 pages.
https://doi.org/10.1145/3321513

1 INTRODUCTION

Nowadays, stereo displays have been prevalently deployed in people's life. We may find their usage in VR, 3D movies, 3D televisions, and the like. However, the diversity of resolutions of stereo display screens has made it impractical to provide data of different sizes for different screens, which gives rise to a bottleneck for the popularization of stereo display technologies. Various algorithms have been designed to adjust the resolutions of the stereo images dynamically according to the resolutions [1, 4, 19, 20, 26, 30, 36, 37]. Each pair of stereo images consists of two correlated images.

The work was supported by National Science Foundation Grant of China 61370160, 61772149, Guangdong Province Natural Science Foundation Project (2015A030313578), Guangdong Scientific and Technological Plan Project 2015B010106005, 01904010228.
Authors' addresses: Y. Gong, Guangdong University of Foreign Studies, Guangzhou Higher Education Mega Center, Panyu District, Guangzhou, China, 510006; email: gongyongyi@gdufs.edu.cn; S. Li and K. Wattanachote, Guangdong University of Foreign Studies, Guangzhou, China; emails: lsr97mail@yeah.net, kanoxak@hotmail.com; X. Luo, Guilin University of Electronic Technology, 541004, Guilin, China; email: luoxn@guet.edu.cn.

1551-6857/2019/08-ART69 $15.00
https://doi.org/10.1145/3321513

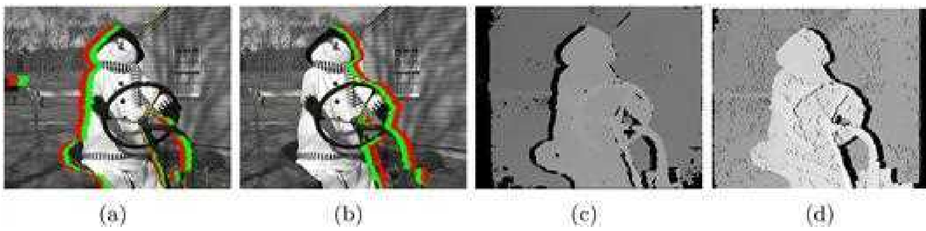


Fig. 1. Green areas on (a) and (b) were occluded red areas on each other image. (c) was the disparity map calculated by SGM [11], and (d) was the disparity map calculated for the right image based on Algorithm 2 introduced in Section 3. Both of the disparity maps were necessary, as monocularly visible pixels on both images should be equivalently considered. A detailed explanation is carried out in this section.

The key point in stereo image retargeting is to maintain the correspondence between pixels from both images in the stereo pair, whilst depth information should be preserved. At the same time, the distortion of high-saliency regions is to be avoided as much as possible in order to maintain an appealing visual experience for viewers after retargeting.

According to the correspondence between pixels on both sides, pixels in stereo images are classified into four categories: ordinary binocularly visible pixels from both sights, occluded pixels, occluding pixels, and out-of-sight pixels from either perspective which are monocularly visible. The first category exists in pairs on the stereo images, of which the modifier "ordinary" is meant to exclude the third category, a.k.a. occluding pixels. They are correspondent according to the disparity, while occluded pixels, and monocularly out-of-sight pixels, only exist on one of the images in the stereo image pair. Figure 1 gives a visual example. As we will explore in Section 3, there are no counterparts for them on the other image. Stereo retargeting algorithms ought to preserve the relationship of correspondence of four categories of pixels stated above.

We noticed that although people consider occlusion when designing stereo retargeting algorithms, a problem is often neglected in current methods according to literature we reviewed. They only chose the left image as the reference image, without considering occluding and occluded regions on the right image. In fact, occluded regions on the right image have corresponding regions on the left image, which will cause the violation of geometric consistency when deleted, and thus the declination of imaging effects and visual comfort.

Occluded pixels on the left image are considered in the method reported by Basha et al. [4]. But there are also monocularly visible pixels on the right image. Each of these pixels has its corresponding binocularly visible counterpart, or an occluding pixel. If we remove stereo seams containing these occluding pixels, occluded pixels on the right image will remain solitary without occluding pixels accounting for them. If we simply calculate the disparity map based on the left image, this important information will be lost. Thus, it will inevitably break the geometric consistency, so a detection based on the right image is necessary.

Thus, we will emphasize the relationship between occluded and occluding regions. We propose a method of detecting occlusion to solve the problem of avoiding occluded and occluding regions on both of the images in a stereo pair, after which a new scheme of stereo seams selection is introduced. Based on this scheme, we propose an advanced stereo retargeting method of seam carving. We also propose the utilization of edge and line segment detection to better preserve the appearance of the images.

2 RELATED WORKS

Compared to traditional two-dimensional images, stereo images have pixels correspondent to each other on left and right images. With the also thereby produced depth information, these have become a significant challenge waiting to be confronted when researching content-aware stereo image retargeting. Current progress in this field [4, 22, 24, 30, 39] can be separated into discrete and continuous categories.

Discrete methods of stereo image retargeting is based mainly on seam carving. Utsugi et al. [30] applied as a preliminary method the seam carving algorithm to stereo images, which found a corresponding pixel on the right image after finding a pixel to be deleted on the left image. In this case, the deletion was performed on both images. If occluded regions were crossed by a seam on the left image, the part of the seam crossing the region would have corresponding pixels found on the right image. The method made up corresponding pixels on occluded regions of the right image, making it always possible to find legal seams on the right image. However, the strategy of making up corresponding points violated the relationship of occlusion of the images, and severely damaged depth information. A new seam carving algorithm was proposed by Basha et al. [4] later on. It not only considered the correspondence between pixels on both sides, but also took occluding and occluded pixels into account. Their method intentionally avoided these pixels. However, they only avoided these regions on a single image, damaging depth information of images inevitably, and influencing the geometric consistency as well. Fang et al. [8] combined stereo saliency, edge reflection, and energy cost into an energy function, leading to an advance in the effect of retargeting. Basha et al.'s methods have been leveraged in several researches, and discrete methods continue to be developed [3, 5, 10, 27, 36].

Continuous methods were mainly based on mesh distortion. They could be found in various literature [7, 17, 25, 28, 29]. We paid our attention to discrete methods.

Recently, Shao et al. [27] proposed an ameliorated stereo seam carving method, and incorporated stereoscopic visual attention and binocular just-noticeable difference models into the energy optimization, better preserving salient objects and geometric consistency. As for the processing of occluded regions, Shao et al.'s method sought optimal seams on the left image first, and replaced all parts which passed occluded regions, and eventually found matching seams on the right images. Thus, this method was merely avoiding occluded regions on one side. Lei et al. [18] also noticed this. They mainly introduced matching map and importance map scaling factors to cope with seam coupling errors caused by accurate disparity matching. Visual effects were improved. However, occluded regions were still avoided on one side only, and depth information was damaged due to occlusion. The extraction of saliency regions is also a significant topic when retargeting stereo images [32, 33]. A recent method proposed by Wang et al. [34] concerned the production of thumbnails based on stereo saliency detection. With a preferable visual effect after retargeting, occlusion was not considered in the method.

Seam carving is one of the pervasively functioning discrete methods. It provides content-aware image retargeting, reducing or expanding only areas considered to be visually unimportant, leaving salient areas intact, while the stereo image seam carving has further challenges. Objects are projected to both left and right images, and their relationship of correspondence has to be taken into consideration, of which occlusion is an important topic. Existing stereo seam carving methods extend seam carving in different ways. Some avoid occluding areas [4], some improve the energy function [8], and some ameliorate the matching process [18] as described above. The main motivation of our approach, however, would be to fix a defect that results from asymmetric consideration of left and right images in traditional seam carving methods, and include further optimizations such as line segment detection, for that preserving edges of salient objects would also improve visual saliency.

Generally, discrete and continuous stereo retargeting algorithms both have their pros and cons. Discrete stereo retargeting preserves the depth information of images better, though prone to distortion of objects; while continuous methods accentuate the protection of the shape of the object.

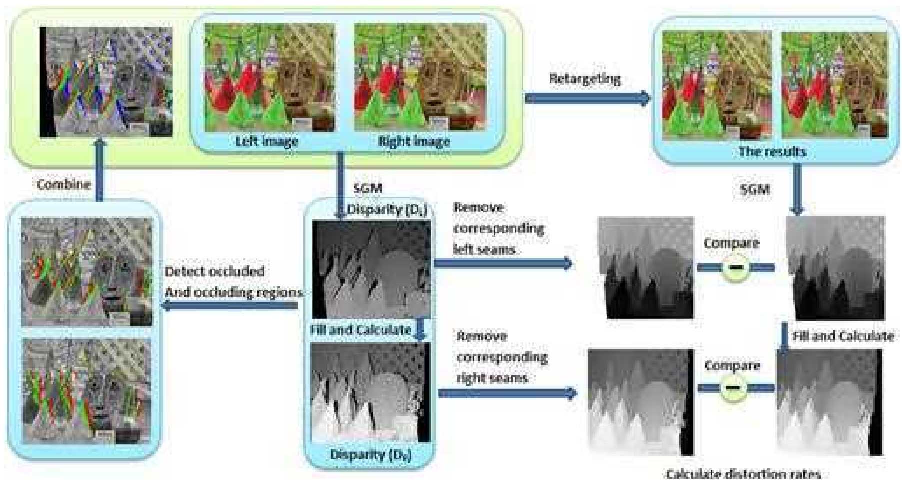


Fig. 2. Our general scheme. The calculation of distortion rates will be covered in detail in Section 5. A disparity map was calculated based on the left image. According to the left image, a disparity map based on the right image was calculated. Because of rounding issues, some gaps appeared and were filled with nearby values. After that, occluding and occluded regions were detected, and combined onto a single image for the convenience of calculation. Stereo retargeting based on dynamic programming was performed to give the results. Furthermore, corresponding seams removed during retargeting were also removed in D_L and D_R, attaining a pair of reduced images. Then we performed disparity detection on the results to get another pair of images. These two pairs will be compared for distortion rate assessment in Section 5.

Nevertheless, current methods have not completely preserved geometric consistency rigorously as some of them claimed to have. Occluding and occluded pixels might be deleted, which caused the violation of the relationship of correspondence among pixels, broke the occluded-occluding relations in the original scene, and influenced the visual effects.

3 OCCLUDED AND OCCLUDING REGIONS IN STEREO IMAGES

According to the principles of stereo vision, occluded and occluding pixels abound in both of the images in a stereo image pair. Through evidence discovered in the research of psychophysics [23], occlusion is considered as a significant factor when the human eyes are observing objects in three-dimensional space. It is necessary to take into account occluded and occluding regions in both of the images of the stereo pair when retargeting stereo images as in Figure 2.

3.1 The Relationship between Occlusion and the Human Visual System

Our new strategy considering occlusions on both side of the stereo images aims at producing better visual comfort. And the human visual system is closely related to its perception of the depths of objects. Depth difference among objects, which is larger than the just noticeable depth difference, will be sensitively detected by the human visual system [15].

The destruction of geometric consistency will lead to the distortion of depth information, which is essentially breaking the occlusion relationship. Imagine that an object is mistakenly deleted from monocular sight, depth information or occlusion relationship around this object will suffer from distortion.

We consider a simple example. According to the definition of geometric consistency of Basha et al. [4], pixels in the blue areas on Figure 11(g) should not be removed during stereo retargeting, because if they are removed, corresponding areas on the right images will be removed, and the

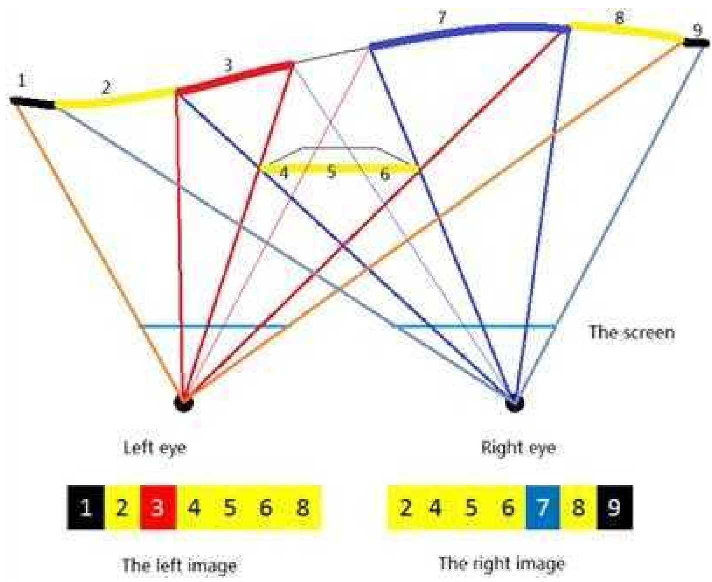


Fig. 3. This picture illustrates the relationship between objects in actual scene and their appearance on stereo image pair.

areas they occluded before become no longer occluded but still remain monocularly visible. This would apparently cause distortion of depth information in cases like Diana's shoulders of Figure 11 and Figure 12, correspondingly influencing visual comfort based on the perception through human visual system.

3.2 The Reason of Considering Occlusion on Both Sides

Discrete image retargeting on 3D image pairs are a little bit more complicated than retargeting on a single image. When we select an 8-connected vertical seam, with a pixel on every line, on one image of a pair of 3D images, we need to find corresponding pixels on the right images so that we are deleting the same contents on both images. Thus, we need to make use of the SGM [11] algorithm as done in Basha et al.'s method [4] to determine the correspondence of pixels between two images.

Retargeting of a pair of images faces inevitably the problem of occlusion. As demonstrated in Figure 3, we can see that, in the actual scene, some of the objects can be seen only on one of the screens, but not on the other. These monocularly visible areas include two types. One type is monocularly out-of-sight areas. As demonstrated in Area 1, Figure 3, the black area is too far on the left to appear on the right image. The other type is occluded areas as demonstrated by Area 3 on Figure 3 as an example. The object in the middle of the scene blocked it, thus it could not be seen on the right image.

The areas on both of the images, or the binocularly visible areas, are marked as yellow on Figure 3. Each of the pixels in these areas have a counterpart on the other image. But some of the pixels in binocularly visible areas are special in the sense that they occlude other pixels. For example, Areas 4 and 5 on Figure 3 are the reason why Area 3 (red area) is not visible on the right image. Hence, for the left image, we name Areas 4 and 5 as occluding areas.

We can easily see that pixels in the occluded areas, like Area 7 (blue area), have corresponding pixels that occludes them, which we call *occluding pixels*. Occluding pixels, like those in Areas 5 and 6, appear on both images, while occluded pixels appear only on a single image like those in Area 7.

During the retargeting process, if we delete occluded pixels on a single image like those in the red areas of Figure 3, we will not be able to find its counterpart on the other image. So the deletion of occluded pixels should be avoided. If we delete occluding pixels in pairs like those in Areas 4, 5 and 6 of Figure 3, the corresponding occluded pixels, those of the red areas and blue areas, will lose their covering, or in other words, can no longer find the occluding pixels correctly. So the deletion of occluding and occluded areas, like the yellow, red, and blue areas in Figure 3, should be avoided during retargeting, since geometric consistency will be violated without doing so.

Former researchers like Basha et al. [4] had noticed this problem. They defined these areas as Equation (1). They chose the left image as the reference image, calculating the occluding and occluded areas on the left image (the yellow and red areas in Figure 3), and avoided deleting them when selecting seams. This method seems to have preserved geometric consistency. However, one important point was missed. What if seams we selected include occluding pixels on the left image? As illustrated above, pixels in Areas 5 and 6 are occluding pixels corresponding to occluded pixels in Area 7. With a traditional method, Area 5 is avoided but Area 4 is still subject to deletion. Therefore, if we delete pixels in Area 4, occluded pixels in the blue area will lose their covering, or cannot find any corresponding occluding pixels any more. This violation of geometric consistency occurs since traditional methods do not consider occluded regions on the right image, which is why we should consider occlusion on both sides.

3.3 Occluded and Occluding Regions

Occluded regions are spatially coherent groups of pixels that can be seen in one image of a stereo pair but not in the other. In the three-dimensional scene of binocular stereo retargeting, if a binocularly visible object occludes another object so that the latter object is merely monocularly visible, we claim that the former object is the occluding object, while the latter one is the occluded object. Occluding and occluded pixels are correspondent to each other due to the disparity.

We give the definition of occluding and occluded pixels. The notation of $P_L(i,j)$ is used to designate the pixel on the ith column and the jth row of the left image. Similarly, we designate the corresponding pixel on the right image with $P_R(i,j)$.

Definition 1. Suppose $I = \{I_L, I_R\}$ is a pair of stereo images with the size of $m \times n$, and D_L is the disparity map with the left image I_L as the reference image. In the $j(1 \le j \le m)$th row of I_L, if there exists a pixel $P_L(i_1,j)$, for all $s(s \neq 1)$ pixels $P_L(i_1,j), P_L(i_2,j), \ldots, P_L(i_s,j)$ satisfying

$$i_k + D_L(i_k,j) = i_1 + D_L(i_1,j),\ k = 2,\ldots,s, \tag{1}$$

where

$$|D_L(i_1,j)| > |D_L(i_k,j)| \tag{2}$$

We claim that $P_L(i_2,j),\ldots,P_L(i_s,j)$ are **occluded pixels on the left image**, while $P_L(i_1,j)$ are their corresponding **occluding pixels on the left image**.

We claim that the set of occluded pixels on the left/right image is the **occluded regions** on the left/right image, while the set of their corresponding occluding pixels are **occluding regions**.

3.4 Detecting Occluded and Occluding Regions on Both Images

In this section, we propose two algorithms to detect occluded and occluding regions on the left and right images, respectively.

To consider occlusion on both sides, we need further calculation to determine occluded and occluding areas on the right image (Areas 5 and 6 and the blue area on Figure 3), which means that we should apply (1) on the right image. Hence, we need a disparity map based on the right image (D_R). Traditional usage of SGM [11] only yield disparity map based on the left image (D_L).

ALGORITHM 1: The detection of occluded and occluding regions on the left image

Input: D_L, I_L
Output: $U_{Occluding_L}, U_{Occluded_L}$
Step 1. Initializations. Define $[m, n] = sizeof(D_L)$, $D_{map} = zeros(m, n)$, $l = (1, 2, \ldots, n)$ and create a matrix D_l with the size of $m \times n$ which has every row the same as l;
Step 2. $D_{map} = D_l + D_L$;
Step 3. On the jth row in D_{map}, for all $s(s \neq 1)$ pixels $P_L(i_1, j), P_L(i_2, j), \ldots, P_L(i_s, j)$ satisfying $D_{map}(i_1, j) = D_{map}(i_k, j), k = 2, \ldots, s$, if we have $|D_L(i_1, j)| > |D_L(i_k, j)|$, then $P_L(i_2, j), \ldots, P_L(i_s, j) \in U_{Occluded_L}$, $P_L(i_1, j) \in Occluding_L$

The trick of mirroring the image pair to get the disparity map based on the right image does not work well. Hence, we need another process, which we call "Fill and Calculate".

In this process, we try to calculate D_R based on D_L. For binocularly visible pixels, the act of merely negating the disparity values and inserting them in the corresponding position may suffice. For the gaps between them, we fill them with nearby values. Detailed illustrations of "Fill and Calculate" are in Algorithm 2.

After getting D_R, we are able to apply (1) on the right image and find out occluding and occluded areas (Areas 5 and 6 and the blue area) on the right image. For convenience, when we retarget the image pair, we only mark areas needed to be avoided on the left image. To avoid the deletion of Areas 4 and 5, we need to mark them on the left image based on their position on the right image, and D_R. Detailed illustrations are shown in Algorithm 2.

During retargeting, we choose seams on the left image to avoid occluding and occluded areas (Areas 4 and 5 and the red area on Figure 3), and corresponding areas of occluding areas on the right image (Areas 5 and 6 on Figure 3), which traditional methods have not considered.

First, we propose Algorithm 1 to detect occluded and occluding regions on the left image, which can be done according to Definition 1 as follows:

We make use of some currently mature stereo matching algorithm (we use SGM [12]), acquiring D_L, the disparity map based on the left image. And according to the definition of occluded and occluding pixels, we acquire $U_{Occluded_L}$ which are occluded regions on the left image, and their corresponding occluding regions $U_{Occluding_L}$. Algorithm 1 illustrates the details. Next, we state how to detect occluded regions and corresponding occluding regions on the right image and propose Algorithm 2. Suppose D_L is the disparity map based on the left image, we define a disparity information matrix based on the right image as follows:

$$k + D_L(k, j) = i \wedge P_L(k, j) \notin U_{Occluded_L} \tag{3a}$$

$$D_R(i, j) = \begin{cases} -D_L(k, j), & if\,(3a) \\ Null, & if\, not\ (3a) \end{cases} \tag{3b}$$

We mark corresponding disparity on the disparity information matrix according to disparity of binocularly visible regions on the left image. Now there is still some undefined regions on D_R, including monocularly out-of-sight regions and occluded regions on the right image.

If $D_R(i, j)$ is *Null*, then $P_R(i, j)$ belongs to U_{Null}. If pixel $P_R(i, j)$ belongs to U_{Null}, monocularly out-of-sight regions are

$$U_{OutOfSight_R} = \{P_R(i, j) | P_R(k, j) \notin U_{Null},\ \forall k < i\}. \tag{4}$$

Apart from monocularly out-of-sight regions, occluded regions on the right image are

$$U_{Occluded_R} = U_{Null} - U_{OutOfSight_R}. \tag{5}$$

ALGORITHM 2: The detection of occluded and occluding regions on the right image

Input: I_L, I_R
Output: $U_{Occluding_L}, U_{Occluded_L}, U_{Occluding_R}, U_{Occluded_R}$
Step 1. Calculate disparity map D_L according to I_L with currently existent stereo disparity calculating methods;
Step 2. Detect occluded and occluding regions on the left image with Algorithm 1;
Step 3. With disparity information of binocularly visible regions on D_L, we define disparity information matrix on the right image D_R, and if $D_L(i,j) \notin U_{Occluded_L}$, we have $D_R(i + D_L(i,j), j) = -D_L(i,j)$;
// The following two steps are "Fill and Calculate" in Figure 2;
Step 4. Find monocularly out-of-sight regions on the right image, and for the jth row on the right image;
for $i = 1$ *to* n **do**
  if $D_L(i,j) \notin U_{Occluded_L}$ **then**
    $D_R(i + D_L(i,j), j) = -D_L(i,j)$;
  end
end
Step 5. For undefined regions apart from monocularly out-of-sight regions, which are occluded regions on the right image, we fill them with values on the nearest valid pixels;
Step 6. Similar to Algorithm 1, we use D_R and I_R as our input, acquiring occluding regions $U_{Occluding_R}$ and occluded regions $U_{Occluded_R}$ on the right image;

For pixel (i,j) in $U_{Occluded_R}$, we make it up with the smaller disparity of the two nearest defined pixels from two directions in the same row.

$$D_R(i,j) = min\{D_R(e_l), D_R(e_r)\}, P_R(i,j) \in U_{Null} \tag{6a}$$

in which

$$e_l = (max\{k|k < i \wedge k \notin U_{Null}\}, j) \tag{6b}$$

$$e_r = (min\{k|k > i \wedge k \notin U_{Null}\}, j) \tag{6c}$$

Similar to Definition 1, we can find corresponding occluding regions $U_{Occluding_R}$, and occluded regions $U_{Occluded_R}$ on the right image. Algorithm 2 illustrates the details.

3.5 Scheme of Avoiding Occlusion on Both Sides

According to the analysis above, we have known that occlusion exists on both sides of the stereo pair. Thus, it is insufficient to avoid only occluded and occluding regions in one of the images in the stereo pair. We propose an advanced scheme considering occluded and occluding regions on both sides during stereo image retargeting, which is not considered in current stereo retargeting algorithms, causing the violation of geometric consistency and the declination of visual comfort.

The corresponding relationship between occluded and occluding pixels is a significant issue regarding stereo retargeting. For each occluded pixel, which appears only on one side of the stereo pair, there must be a binocularly visible pixel causing the occlusion of it, which appears in pairs on both sides of the stereo pair. Thus, there are three pixels forming the relationship, within which any one deleted will cause the destruction of geometric consistency.

If the correspondence between occluded and occluding pixels on the left and right image is neglected, pixels of the blue regions on Figure 4 may be chosen as part of the seams and thus deleted. This may cause the deletion of pixels of occluding regions on the right image. If such a pair of pixel is deleted, the corresponding occluded pixel will no longer be occluded, but it does not have a counterpart on the other image. This apparently violates the definition of geometric consistency defined by Basha et al. [4]. The left image is designated as the reference image when

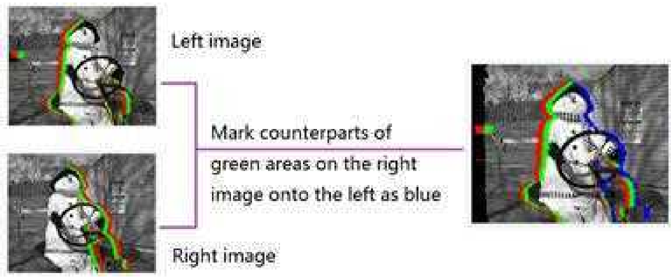


Fig. 4. Red and green parts on the left are occluded and occluding regions in left and right images, respectively. Blue parts are corresponding regions of occluding regions on the right image onto the left image. Apparently, the deletion of pixels in the blue regions will violate geometric consistency of the images. Former method like Basha et al.'s [4] did not consider blue regions, so they broke geometric consistency to some extent.

we select seams. We select seams on the left image, and find the correspondent seam on the right image. Though regions we need to avoid exists on both of the images, we can mark regions needed to be avoided only on the left image as long as we reflect occluding regions on the right image onto the left image. In other words, we need to reflect $U_{Occluding_R}$ onto the left image, on which we acquire corresponding regions of the right image.

$$U_{OccludingRonL}(i + D_R(i, j), j) = U_{Occluding_R}(i, j). \tag{7}$$

When seams are selected, we make use of an avoiding scheme on both sides.

Scheme 1. *In order to protect geometric consistency of stereo images, we should avoid out-of-sight regions, occluding and occluded regions simultaneously when stereo retargeting is performed. Specifically, out-of-sight regions, occluding and occluded regions on the left image, and corresponding regions of occluding regions on the right image.*

During the retargeting process, we choose seams on the left image with the least energy cost, which is calculated through the energy function described in detail in the following section to avoid areas stated above. We find the corresponding seam on the right image as well, since the occluding regions on the right image have already been marked on the left image.

4 ENERGY FUNCTION

In this section, the calculation of energy function is discussed in our new content-aware stereo retargeting algorithm, which is based on a strategy avoiding occluded and occluding regions, in detail. Using this energy function, we are able to determine the energy cost of adding a single pixel onto or removing a single pixel from the stereo pair. Together with the areas needed to be avoided in Section 3, we adopt a dynamic programming method introduced by Basha et al. [4], by adding or deleting stereo seams in turn.

Let $I = \{I_L, I_R\}$ be a pair of stereo images of the size $m \times n$, we build an energy function E_{total} based on the left image, by designating its values on pixel (r, c) consisting of $E_{intensity}$, E_{3D}, and $E_{attention}$.

$$E_{total}(r, c, c^{\pm}) = E_{intensity0}(r, c, c^{\pm}) + \alpha E_{3D}(r, c, c^{\pm}) + \beta E_{attention}(r, c). \tag{8}$$

We use α and β as the experimental parameters measuring the importance of corresponding terms. When the importance of disparity exceeds that of intensity, we designate a higher value for α. Through experiments, we select 0.5 and 0.5, the same value respectively for α and β, while other values may also be considered. $c^{\pm}$ stands for the column coordinate of the $(r-1)$th pixels. Since the

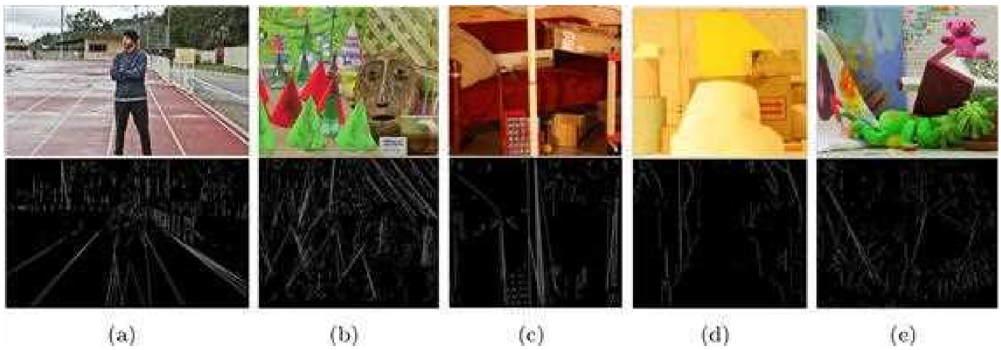


Fig. 5. The images on the first row were the original images. And the images on the second row were the result of applying Line Segment Detection [31]. Line segments with a slope angle lower than 30 degrees were removed for the convenience of vertical seams.

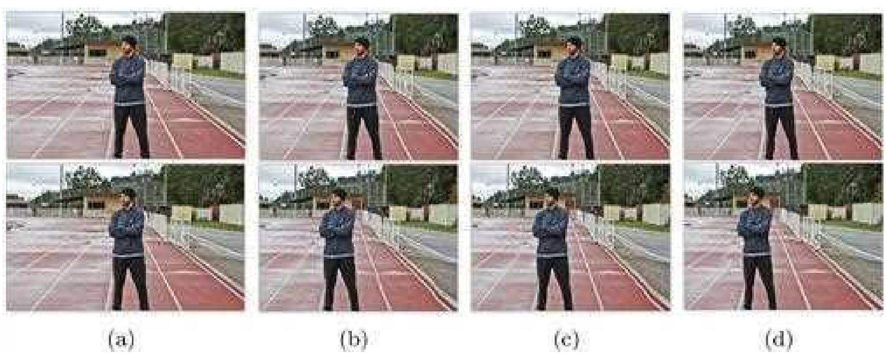


Fig. 6. Different choices of ϵ. (a) were original images, (b) were our results with ϵ set to 0.5. (c) were our results with ϵ set to 1. (d) were our results with ϵ set to 10. All photos were shrunk by 17%.

pixels of a vertical seam are not necessarily continuous, $c^{\pm}$ can be of any column in the $(r-1)$th row.

4.1 Energy of Attention

We propose the energy of attention in order to emphasize the sensitivity of the human eyes about contours, fringes and linear features of the images. If fringes of the images, especially linear features, are damaged, the distortion of the image may be inevitable. To meet with this situation, $E_{attention}(r,c)$ is defined below.

$$E_{attention}(r,c) = \delta \cdot \alpha_{Affected}(r,c) + \epsilon \cdot \alpha_{Line}(r,c) \tag{9}$$

We mark the detected line segments with LSD [31] on matrix α_{Line}, originally with the same size as the left image. Seams crossing the detected lines will have higher energy so that they are less likely to be chosen. In some cases like that of Figure 5, there are so many lines that crossing is inevitable. For this situation, we propose a strategy of avoiding crossing the same line in the same position for many times, thereby preserving visual comfort to some extent.

At the beginning of retargeting, α_{Line} is marked as 1 where line segments exist and 0 otherwise. As a seam is selected and deleted from the left image, it is also deleted on α_{Line}. If the seam crossed a line at some point, the pixels near the point will have their value of α_{Line} increased, so that following seams are less likely to cross them.

ALGORITHM 3: Avoiding seams getting too close to each other

Input: Deleted pixel (i, j), matrix $\alpha_{Affected}$ initialized to all zeros
Output: Matrix $\alpha_{Affected}$
for all i', j' such that $|i - i'| <= 2$ and $|j - j'| <= 2$;
  if $\alpha_{Affected}(i', j') == 0$, then $\alpha_{Affected}(i', j') = 1$;
  else if $\alpha_{Affected}(i', j') > 0$, then $\alpha_{Affected}(i', j') = \alpha_{Affected}(i', j') + \alpha_{Affected}(i', j')^{-1}$;

ALGORITHM 4: Avoiding seams crossing a line segment at the same point too many times

Input: Deleted pixel (i, j), matrix α_{Line} on which line segment areas are marked as 1 and other areas are marked as 0
Output: Matrix α_{Line}
for all i', j' such that $|i - i'| <= 2$ and $|j - j'| <= 2$;
  if $\alpha_{Line}(i', j') == 0$, then $\alpha_{Line}(i', j') = 0$; //No change
  else if $\alpha_{Line}(i', j') > 0$, then $\alpha_{line}(i', j') = \alpha_{line}(i', j') + 1$;

A similar strategy is used in seam deletion as well. When a seam is deleted, we increase nearby $\alpha_{Affected}(r, c)$ values, so that seams do not get too closed to each other. It is another way to improve visual comfort.

The use of $\alpha_{Affected}(r, c)$ accounts for the distortion caused by seams which are selected too closed to each other. In our experiments, when a pixel (i, j) is deleted as a part of a seam, the pixels around it within the radius of 2 are more suitable than others through experiments. Thus, we introduce Algorithm 3 to avoid visual distortion.

Similarly, for α_{Line}, we have Algorithm 4 in which α_{Line} is used to mark the position of line segments onto a matrix. Every time a seam crosses a line at some point, the value of α_{Line} around that point is increased so that a line is not crossed twice at the same point by different seams.

The values for δ, ϵ should apparently influence E_D. In our experiments, we chose values 0.01 and 2, respectively, with other suitable values also available. Figure 6 illustrates the influence of different choices of ϵ. Next, we discuss the $\alpha_{Line}(r, c)$.

The human eyes are not only sensitive about salient regions, but also linear features of the images. If there exists too much distortion in lines on the images, the visual comfort of images can be severely influenced. $\alpha_{Line}(r, c)$ is defined so that seams are not crossing line segments in the same place.

With Line Segment Detection [31], we get our matrix of line segments. We integrated line segments with similar slopes and approached end points into a single line segment to improve visual comfort of the contour. For pixels which were designated as part of a line segment, we have $\alpha_{Line}(r, c) = 1$.

Some lines detected with LSD [31] are closed to each other and with very similar slopes. They usually appear to the human eyes as a single line segment. Thus, we need to connect these lines, or seams will cross these gaps causing visual discomfort.

For all the line segments detected with LSD [31], if end points of two lines are lower than a set threshold, and the slopes are closed enough, or in other words, the difference is lower than a threshold, we add a long line covering these two lines to the line set.

For two end points (x_1, y_1), (x_2, y_2) of a line segment l in any detected set of line segments S_l, and for pixel (i, j), if we have $y_1 \leq j \leq y_2$ and

$$i' = (j - y_1) \cdot \frac{x_2 - x_1}{y_2 - y_1} + x_1, |i' - i| < t_1, \tag{10}$$

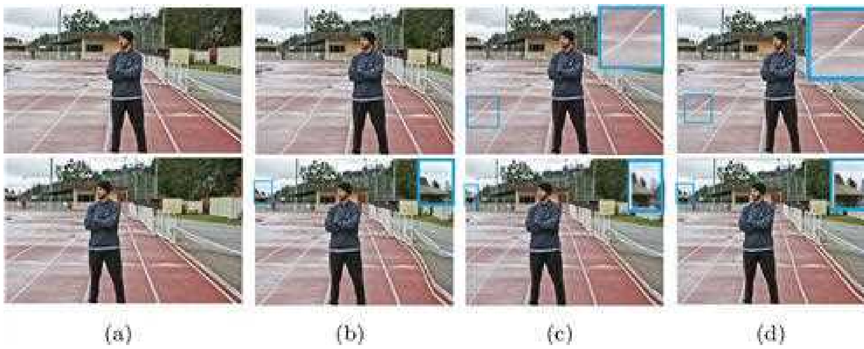


Fig. 7. The images on the first row are left images and the images on the second row are right images. Column (a) illustrates the original images of Man dataset. Column (b) demonstrates the output results of Basha et al.'s method. Column (c) demonstrates the results of Li et al. [20]. Column (d) illustrates the results of our method, on which line segment detection has been applied on the image retargeting process, noticed by the straight lines of the runway. As shown at the bottom right corner of our right image in column (d), some distortion is inevitable as part of the image belonging to monocularly out-of-sight regions. The results of Basha et al.'s were obtained by their codes posted on the Internet. Though not the same as the results they claimed to have [4], other researchers, such as Lin et al. [21], have also obtained the same results similar to ours using Basha et al.'s codes.

then (i, j) belongs to this line segment. We have $t_1 = 2$ due to our experiments, other favorable values are also within consideration.

Then, for any pair of two lines l_1, l_2 with slopes of k_1, k_2 in S_l, if we have

$$(\arctan(k_1) - \arctan(k_2)) \cdot \frac{180}{\pi} < t_2 \tag{11}$$

we integrated l_1, l_2 into one line segment, and selected farthest two points from the four end points of the two line segments as the end points of the new line segment.

Through experiments, we have $t_2 = 5$ with other favorable values also within consideration. Figure 7 demonstrates the result after applying Line Segment Detection [31].

4.2 Energy of Intensity

Deleting a vertical seam will change the neighbourhood between some of the pixels in the stereo image pair. To avoid distortion, we demanded a minimum energy increase after deleting a seam. $E_{intensity}(r, c, c^{\pm})$ is used to measure the difference in intensity between a corresponding pair of pixels in the left and right images on position (r, c).

$$E_{intensity}(r, c, c^{\pm}) = E_L(r, c, c^{\pm}) + E_R(r, c, c^{\pm}). \tag{12}$$

Stereo pixels at position (r, c) are correspondent due to disparity in the image pair. Thus, in Equation (12), E_L, E_R are calculated according to $E(r, c, c^{\pm})$:

$$E(r, c, c^{\pm}) = E^v(r, c, c^{\pm}) + E^h(r, c), \tag{13}$$

where E^v measures vertical changes and E^h measures horizontal ones. Horizontally,

$$E^h(r, c) = |I(r, c + 1) - I(r, c - 1)| \tag{14}$$

Vertically, stereo pixels at $(r - 1, c^{\pm})$ are concerned.

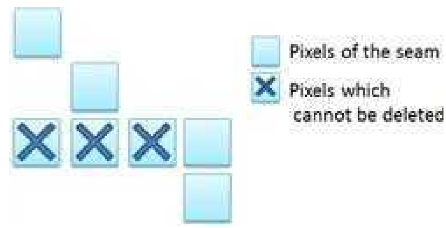


Fig. 8. A "jumping" seam.

$$E^v(r,c,c^{\pm}) = \begin{cases} V_1 + Dist \cdot (c - c^{\pm}), c^{\pm} < c \\ 0, c^{\pm} = c \\ V_2 + Dist \cdot (c^{\pm} - c), c^{\pm} > c, \end{cases} \tag{15a}$$

where $Dist$ is the coefficient for distance. If the distance between c and $c^{\pm}$ is so large that an apparent crack is produced, visual comfort will be influenced.

The seam finding could not always be 8-connected components, since when there are continuous pixels needed to be avoided blocking the way as Figure 8 illustrated, that a seam must "jump" to another position. If two pixels in adjacent lines in the seam are too far away from each other, distortion will be profound. We should shrink $Dist$, the distance between them, as possible. We increase energy cost as when $Dist$ increased. We had some empirical experiments to determine the relationship between the increase of energy and the parameter $Dist$.

The selection of the coefficient for distance should cause apparent influence on E^v. And we determined that the relation should be linear when $Dist$ is small and quadrant when $Dist$ is high.

In our experiments, the calculation below is used.

$$x = |c - c^{\pm}| \tag{15b}$$

$$Dist = \begin{cases} a \cdot x, x < d \\ b + c \cdot x^2, x \geq d \end{cases} \tag{15c}$$

$$V_1 = \sum_{k=c^{\pm}+1}^{c} |I(r-1,k) - I(r,k-1)| \tag{15d}$$

$$V_2 = \sum_{k=c+1}^{c^{\pm}} |I(r-1,k-1) - I(r,k)| \tag{15e}$$

In order to make the influence more apparent when the distance between c and $c^{\pm}$ is smaller to achieve a better effect, we arranged our parameters so that $a = 0.1$, $b = 5$, $c = 0.05$, and $d = 50$ according to our trials and observations.

Besides the linear features on the images, human eyes are also sensitive about the contour of images as shown in Figure 5. Superpixel segmentation is performed on images so that images are zoned, and seams do not tend to horizontally cross the boundaries of the zones. We expanded (13) and the coefficient for superpixel segmentation α_{super} was added.

$$E(r,c,c^{\pm}) = (E^v(r,c,c^{\pm}) + E^h(r,c)) \cdot \alpha_{super}(r,c,c^{\pm}), \tag{16a}$$

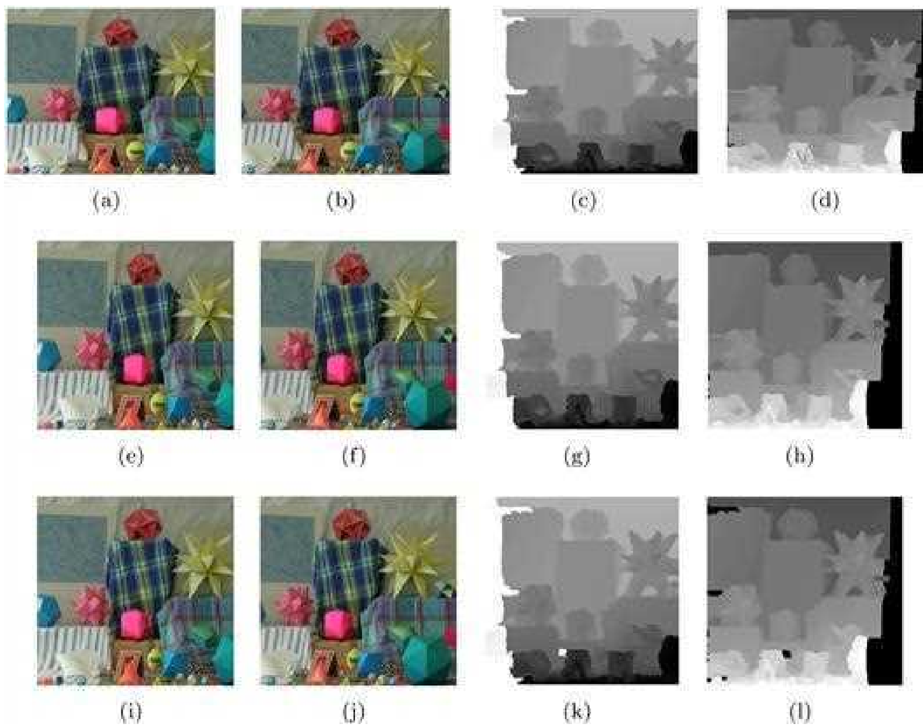


Fig. 9. (a) and (b) were of the original stereo pair of Mob Datasets. While (e) and (f) were our results, (i) and (j) were results of Basha et al.'s method for comparison. Correspondingly, (c) and (d) were disparity maps calculated by SGM based on the left and right images, respectively. (g) and (h) were disparity maps of our results, and (k) and (l) were those of Basha et al.'s results. Disparity maps based on the right image were calculated with the method introduced in Algorithm 2.

where

$$\alpha_{super}(r,c,c^{\pm}) = \begin{cases} \beta_{super}, c^{\pm} \neq c \wedge L_{super}(r-1,c^{\pm}) \neq L_{super}(r,c) \\ \gamma_{super}, c^{\pm} = c \wedge L_{super}(r-1,c^{\pm}) \neq L_{super}(r,c) \\ 1, L_{super}(r-1,c^{\pm}) = L_{super}(r,c), \end{cases} \tag{16b}$$

and $L_{super}(r,c)$ refers to the code for the region to which the pixel on the cth column and rth row belongs. The selection of values of β_{super}, γ_{super} should apparently influence $E(r,c,c^{\pm})$. Through experiments, we select 1.5 and 1.2, respectively. Other favorable values are also within consideration. The cost for a horizontal cross through superpixel boundaries should be larger than that for a vertical cross.

4.3 Energy of Depth

Depth information E_{3D} should be considered when selecting seams on the left image, which is defined as follows:

$$E_{3D}(r,c,c^{\pm}) = E_D(r,c,c^{\pm}) + \beta\,|D_n(r,c)| + \gamma\,|G(r,c)| \tag{17}$$

We choose the values for β *and* γ according to Basha et al.'s correspondent values [4]. E_D measures distortion of the disparity map, which is defined analogous to energy of intensity and

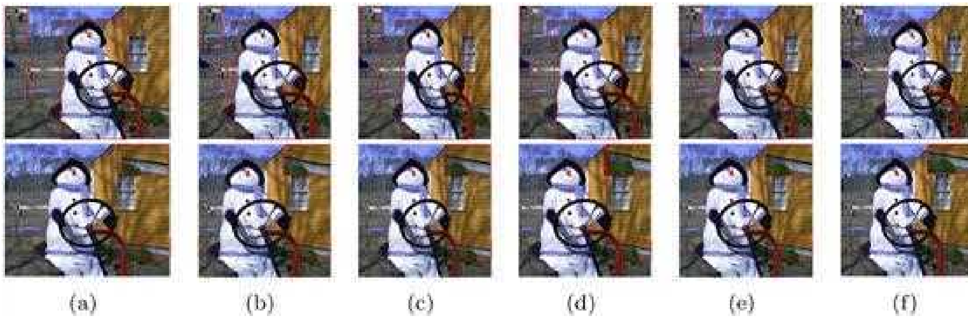


Fig. 10. The first row are left images, and the second row are right images. (a) are the original images and (b)(c)(d)(e) illustrate results from the methods of various researches [2, 4, 18, 20] in turn, respectively. (f) are our proposed method.

Fig. 11. Diana dataset. (a) and (b) were original images, (e) and (f) were our results. (c) and (d) were of Basha et al.'s. All images are shrunk by 17%. (g) marked occluding regions as green, occluded regions as red, and corresponding regions of occluding regions on the right image as blue which traditional stereo retargeting did not consider. Note that, in our proposed method, Diana's left shoulder was preserved with a better visual effect because we considered areas covered in blue.

calculated as formula (12), replacing $E_{intensity}$ as disparity matrix E_D. In actual calculation, D_n must be normalized in advance to minimize influence caused by the difference between ranges of intensity and disparity. A smaller D_n indicates a smaller disparity, which causes less significant influence on stereo effects when pixels were deleted.

Disparity is an important factor affecting visual comfort. An oversized binocular disparity is a significant cause for 3D visual fatigue [6]. $G(r, c)$ measures the difference in intensity between pixels on the left and right images. There are some deviations from reality when SGM is performed, thus when the difference between corresponding pixels on the stereo pair is smaller, the higher certainty that these pixels are geometrically consistent will be reached.

$$G(r, c) = |I_L(r, c) - I_R(r, c + D(r, c))| \tag{18}$$

5 EXPERIMENT RESULTS

As some of the concepts in this research is based on Basha et al.'s method [4], we made use of their code posted on the Internet. We chose data sets Middlebury, Portrait, Flicker, [9, 12, 14] for our experiments. MATLAB was used for implementation, and SGM [11] was used directly to calculate the disparity maps. Figures 9 and 10 compare the results of our method and those of other methods. Figures 13 and 14 focus on some optimizations of local details. Figures 15 and 16 illustrate the evaluation of distortion rates. While Figures 17 and 18 compare the preservation of saliency after applying different methods.

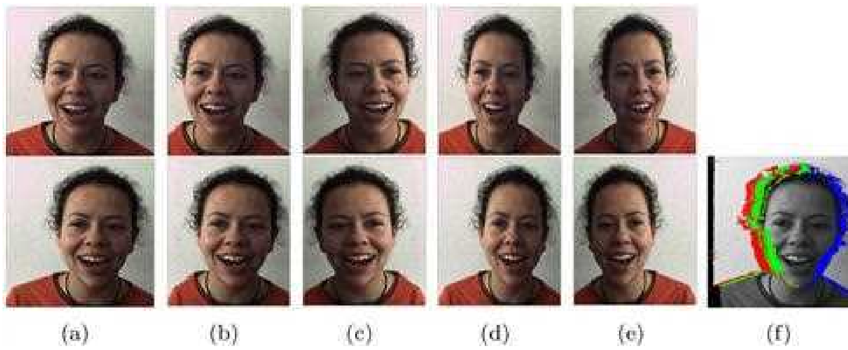


Fig. 12. Another example why we should consider occlusions on both sides. If we flip our original input images and apply seam carving to them, and flip them back, we should have similar results as when we input our images normally without a flip because left and right images are logically symmetric. But traditional seam carving does not preserve occluding areas on the right image, so defects may arise. As illustrated here, (a) were flipped images of Figure 11(a) and Figure 11(b); we put the right image of (a) as the left input and put the left image of it as the right input, getting areas marked on (f) as usual. We apply Basha et al.'s [4] and our method, getting (b) and (d), respectively. We flip them back getting (c) for Basha et al.'s [4] and (e) for ours. If we compare with those of Figure 11, we will find that our proposed method preserved both shoulders regardless of image direction.

5.1 Visual Effects and Distortion Rates

In order to evaluate our results, we compared disparity maps and carried out some quantitative analysis. Binocular distortion is a significant factor influencing visual comfort of 3D images [16], thus analysis of distortion after retargeting should shed some light on visual comfort.

We applied an evaluation of distortion rate similar to Basha et al. [4]. When we performed retargeting on the images, we deleted a seam from the left image at a time. The same seam was deleted on the disparity map based on the left image. At last, we got a smaller disparity map $\hat{D}_o(i,j)$ with the same size as an output. We performed SGM again on the image pair as another output to get another disparity map $\hat{D}_{SGM}$ based on the left image. We compared $\hat{D}_o(i,j)$ and $\hat{D}_{SGM}$ based on the rule below to get distortion rate based on the left image. A visible explanation is shown in Figure 2.

In our proposed method, a disparity map based on the right image allows us to view distortion rate from the perspective of the right image. We calculated D_R based on D_L as illustrated above, and deleted the same seams as is deleted on the right image during retargeting to get a smaller disparity map like above. We performed SGM [11] on the image pair, getting a smaller D_L, and applied "Fill and Calculate", getting a smaller D_R. With similar method as above, we were able to get the distortion rate based on the right image.

Disparity maps can show the difference between pixels in the left and right images. In this part, we shrunk our stereo pair horizontally 17%.

We made use of a method of quantitative analysis from Basha et al. [4]. Here, we applied the same seam carving operations on the original disparity map D_L, and got $\hat{D}_o(i,j)$; suppose $\hat{D}_{SGM}(i,j)$ is the disparity map calculated after seam carving the stereo pair. The evaluating indicator for distortion B is obtained by calculating the difference between ideal and actual disparity. They are defined below (19).

$$B = \frac{1}{N} \sum_{(i,j)} T(i,j) \tag{19a}$$

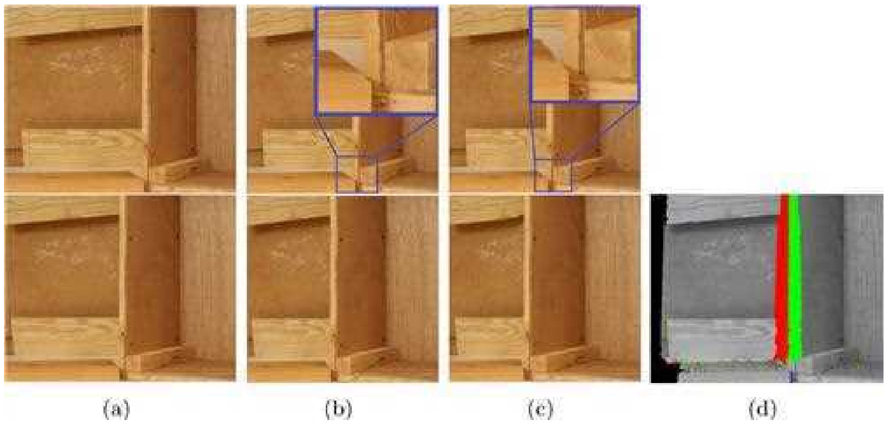


Fig. 13. Wood dataset. (a) were original images, (c) were our results. (b) were of Basha et al.'s. All shrunk by 17%. (d) marked occluding regions as green, occluded regions as red, and corresponding regions of occluding regions on the right image which traditional stereo retargeting does not consider. Note that in our proposed method line segment features were better preserved.

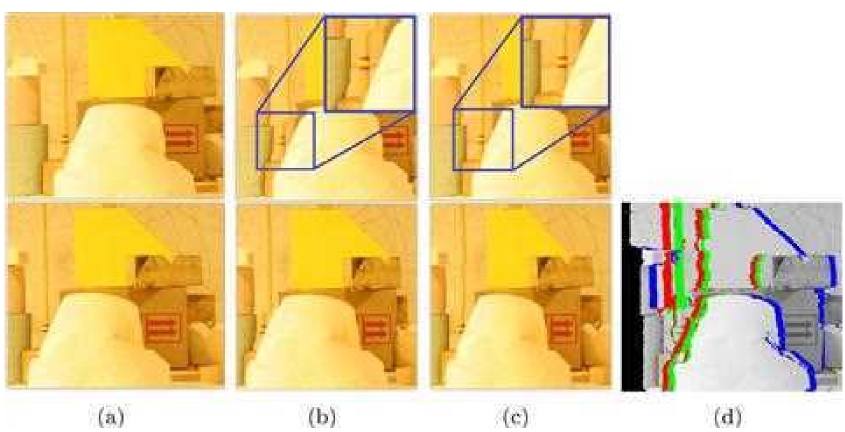


Fig. 14. Lampshade dataset. (a) were original images, (c) were our results. (b) were of Basha et al.'s. All shrunk by 17%. (d) marked occluding regions as green, occluded regions as red, and corresponding regions of occluding regions on the right image which traditional stereo retargeting does not consider. Note that in our proposed method line segment features were better preserved.

$$T(i,j) = \begin{cases} 1, \left|\hat{D}_o(i,j) - \hat{D}_{SGM}(i,j)\right| > 1 \\ 0, \left|\hat{D}_o(i,j) - \hat{D}_{SGM}(i,j)\right| \leq 1 \end{cases} \tag{19b}$$

Formula (19a) refers to the total number of pixels in the image after retargeting, while $T(i,j)$ refers to the stored value at the ith row and the jth column on the image. How many pixels that disparity is changed for more than 1 pixel after retargeting is calculated and a larger value indicates a more severe distortion.

The calculation above merely measures the distortion rate based on D_L. As for D_R, we applied Step 4 in Algorithm 2 on $\hat{D}_o(i,j)$ and $\hat{D}_{SGM}(i,j)$, respectively, to get the counterpart on the right image. We replaced them in formula (19), and calculated the distortion rates once again.

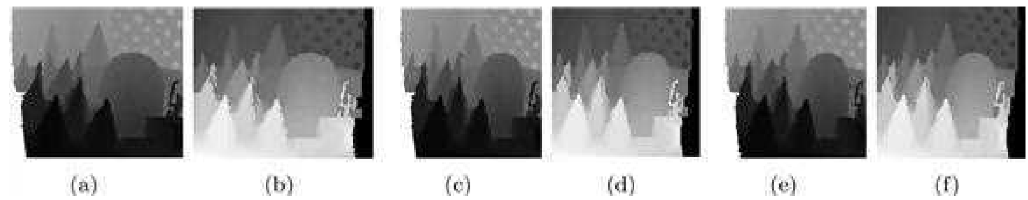


Fig. 15. An example of images used in distortion rate calculation. (a) was D_L, and (b) was D_R acquired by "Fill and Calculate" method in Algorithm 2. And (c) was $\hat{D}_o(i,j)$, and (e) was $\hat{D}_{SGM}$ explained in this section. (d) and (f) were $\hat{D}_o(i,j)$ and $\hat{D}_{SGM}$, respectively, for the right image.

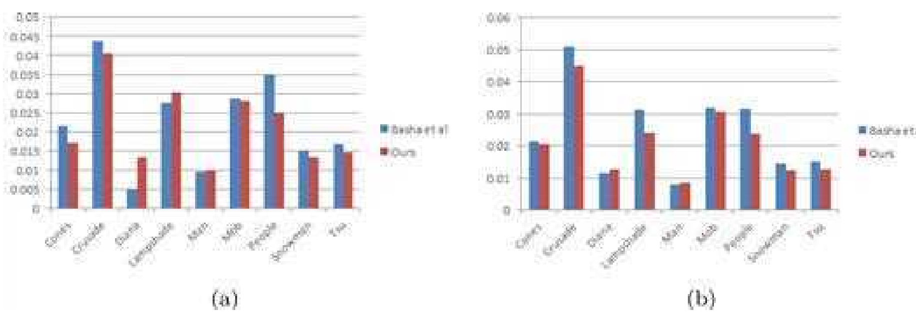


Fig. 16. Distortion rates calculated according to B in (19). (a) was for the left image, and (b) was for the right image. Our methods had lower distortion rate for some of the stereo pairs, especially with regard to D_R.

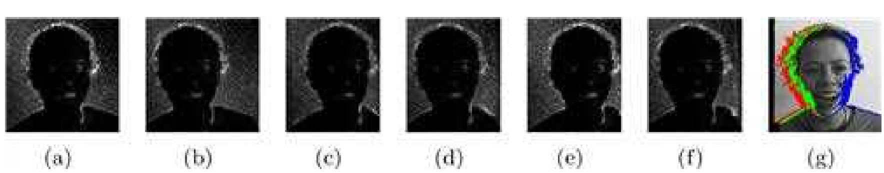
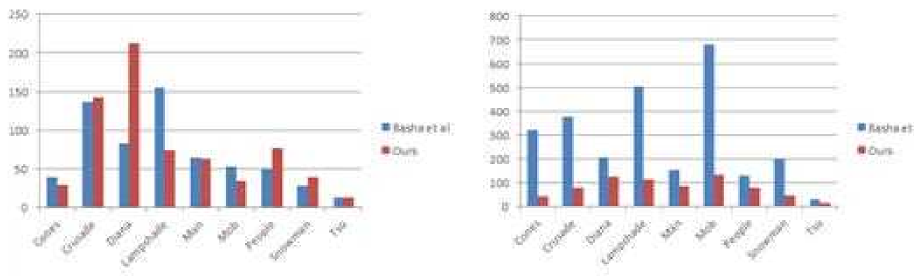


Fig. 17. (a) and (b) were saliency maps calculated for the Diana Datasets. (c) and (d) were those for our results, and (e) and (f) were those for results of Basha et al.'s method. (g) Diana's shoulders were covered with occluded and occluding areas, and corresponding regions of them on the right image.

We can see from Figure 16 that distortion rates of some of our method is higher than those of Basha et al., which is because, in some pictures, regions needed to be avoided are larger in our method. For example, in (g) in Figure 17, regions that are correspondent to occluding regions on the right image (a.k.a. blue areas) on Diana's shoulder make it extremely difficult for a seam to cross consistently, which causes distortion. This is an example of compromising visual comfort for geometric consistency.

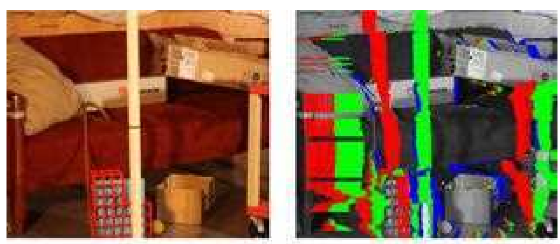
5.2 Visual Saliency Comparison

Visual saliency is detected fast and reliably by the human visual system. In order to further assess our results, we compared the protection of visual saliency in the stereo pair after retargeting in our way with that of Basha et al. First, we calculated the saliency map of the original stereo pair with a method proposed by Hou et al. [13], and applied the seam carving process on it to get $\hat{S}_o$ (just like $\hat{D}_o(i,j)$ in the last experiment). Then, we calculated the saliency map of the retargeted stereo pair to get S_a (just like $\hat{D}_{SGM}(i,j)$ in the last experiment). Now we define accumulative variance



(a) (b)

Fig. 18. Accumulative variance calculated according to (20). (a) was for the left image, and (b) was for the right image.



(a) (b)

Fig. 19. (a) is the original image, and (b) marks the calculated occluded and occluding regions as usual. In this case, regions to be avoided are too large that seams can hardly exist without crossing them. The program will indicate that no more seams can be deleted.

V between these two maps,

$$V = \sum_{(i,j)} (\hat{S}_o - \hat{S}_a)^2 \tag{20}$$

The accumulative variance calculated out of our method and that of Basha et al. As we can see, visual saliency is better preserved in our method, especially on the right image, as we considered occlusions on both sides. As traditional stereo seam carving considers only the left image as the reference image, the distortion on the right image can be profound compared to our method.

6 CONCLUSION

We proposed a scheme that considers occluding and occluded regions of both images of the stereo pair. Our algorithm ameliorates current stereo retargeting algorithms. Although some stereo imaging algorithms have considered occluding and occluded regions, they only cover the circumstances relating to the reference image. Hence, it is insufficient to merely take into account occluding and occluded regions on one side of the stereo pair, which violates the correspondence between pixels during the process of retargeting, and influences visual effects in turn.

We introduced in detail the detection of occluded and occluding regions on both of the images of the stereo pair. We proposed a strategy, upon which we introduced a stereo seam algorithm avoiding occluded and occluding regions. The algorithm completely avoids occluded and occluding regions on both sides, giving a better explication of the three-dimensional scene. Besides, we put in detection of line segments on the images, providing more comfortable visual entertainment for the viewer. Some experimental results have been presented to account for our claims.

There are still some problems in our scheme. For instance, in Figure 19, occluded and occluding regions account for some large continuous proportions of the image, making it difficult for the selection of seams, and gradually causing the distortion of the image. We need further research on these issues. As in the case stated here, we hope to find solutions that can allow us to make our seams cross occluded and occluding regions, though this violates geometric consistency defined above.

Seam carving is originally presented for aesthetic purposes, but it can also be used for malicious purposes. In recent years, seam carving forgery detection has attracted wide attentions [35, 38, 40]. Actually, the investigation of seam carving forgery detection approaches can do benefit to improve seam carving approach. Therefore, we will investigate stereo seam carving detection in future work.

REFERENCES

[1] Shai Avidan and Ariel Shamir. 2007. Seam carving for content-aware image resizing. In *ACM Transactions on Graphics (TOG)*, Vol. 26. ACM, 10.

[2] Bahetiyaer Bare, Ke Li, Bo Yan, Xiaoyu Qi, and Hamid Gharavi. 2015. Pixel fusion based stereo image retargeting. In *IEEE International Conference on Multimedia and Expo (ICME'15)*. IEEE, 1–6.

[3] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan B Goldman. 2009. PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM Trans. Graph.* 28, 3 (2009), 24–1.

[4] Tali Dekel Basha, Yael Moses, and Shai Avidan. 2013. Stereo seam carving a geometrically consistent approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35, 10 (2013), 2513–2525.

[5] Yanxiang Chen, Yifei Pan, Minglong Song, and Meng Wang. 2015. Improved seam carving combining with 3D saliency for image retargeting. *Neurocomputing* 151 (2015), 645–653.

[6] Zaiqing Chen, Junsheng Shi, Xiaoqiao Huang, Lijun Yun, and Yonghan Tai. 2014. Visual comfort modeling for disparity in 3D contents based on Weber–Fechner's Law. *Journal of Display Technology* 10, 12 (2014), 1001–1009.

[7] Sang-Hyun Cho and Hang-Bong Kang. 2013. A new visual comfort-based stereoscopic image retargeting method. In *International Symposium on Visual Computing*. Springer, 290–300.

[8] Yuming Fang, Junle Wang, Yuan Yuan, Jianjun Lei, Weisi Lin, and Patrick Le Callet. 2016. Saliency-based stereoscopic image retargeting. *Information Sciences* 372 (2016), 347–358.

[9] Andrea Fusiello, Emanuele Trucco, and Alessandro Verri. 2000. A compact algorithm for rectification of stereo pairs. *Machine Vision and Applications* 12, 1 (2000), 16–22.

[10] Yongyi Gong, Guan Boliang, and et al. 2016. Stereo seam carving for vertical retargeting. *Chinese Journal of Computers* (2016).

[11] Heiko Hirschmuller. 2008. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30, 2 (2008), 328–341.

[12] Heiko Hirschmuller and Daniel Scharstein. 2007. Evaluation of cost functions for stereo matching. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)*. IEEE, 1–8.

[13] Xiaodi Hou and Liqing Zhang. 2007. Saliency detection: A spectral residual approach. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)*. IEEE, 1–8.

[14] Frédéric Huguet and Frédéric Devernay. 2007. A variational method for scene flow estimation from stereo sequences. In *IEEE 11th International Conference on Computer Vision (ICCV'07)*. IEEE, 1–7.

[15] Seung-Won Jung. 2013. A modified model of the just noticeable depth difference and its application to depth sensation enhancement. *IEEE Transactions on Image Processing* 22, 10 (2013), 3892–3903.

[16] Frank L. Kooi and Alexander Toet. 2004. Visual comfort of binocular and 3D displays. *Displays* 25, 2-3 (2004), 99–108.

[17] Ken-Yi Lee, Cheng-Da Chung, and Yung-Yu Chuang. 2012. Scene warping: Layer-based stereoscopic image resizing. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'12)*. IEEE, 49–56.

[18] Jianjun Lei, Min Wu, Changqing Zhang, Feng Wu, Nam Ling, and Chunping Hou. 2017. Depth-preserving stereo image retargeting based on pixel fusion. *IEEE Transactions on Multimedia* (2017).

[19] Bing Li, Lingyu Duan, Chia-Wen Lin, and Wen Gao. 2014. Region-based depth-preserving stereoscopic image retargeting. In *IEEE International Conference on Image Processing (ICIP'14)*. IEEE, 2903–2907.

[20] Bing Li, Ling-Yu Duan, Chia-Wen Lin, Tiejun Huang, and Wen Gao. 2015. Depth-preserving warping for stereo image retargeting. *IEEE Transactions on Image Processing* 24, 9 (2015), 2811–2826.

[21] Yuqing Lin, Yuzhen Niu, Shuai Zhang, and Junhao Chen. 2017. Anisotropic energy accumulation for stereoscopic image seam carving. In *International Conference on 3D Immersion (IC3D'17)*. IEEE, 1–7.

[22] Yi Liu, Lifeng Sun, and Shiqiang Yang. 2013. Disparity-constrained resolution adaptation for stereoscopic images. In *Pacific-Rim Conference on Multimedia*. Springer, 13–24.

[23] Ken Nakayama and Shinsuke Shimojo. 1990. Da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Research* 30, 11 (1990), 1811–1825.
[24] Jianteng Peng, Jianbing Shen, and Xuelong Li. 2016. High-order energies for stereo segmentation. *IEEE Transactions on Cybernetics* 46, 7 (2016), 1616–1627.
[25] Shaoyu Qi and Jeffrey Ho. 2012. Shift-map based stereo image retargeting with disparity adjustment. In *Asian Conference on Computer Vision*. Springer, 457–469.
[26] Vidya Setlur, Saeko Takagi, Ramesh Raskar, Michael Gleicher, and Bruce Gooch. 2005. Automatic image retargeting. In *Proceedings of the 4th International Conference on Mobile and Ubiquitous Multimedia*. ACM, 59–68.
[27] Feng Shao, Wenchong Lin, Weisi Lin, Gangyi Jiang, Mei Yu, and Randi Fu. 2016. Stereoscopic visual attention guided seam carving for stereoscopic image retargeting. *Journal of Display Technology* 12, 1 (2016), 22–30.
[28] Feng Shao, Wenchong Lin, Weisi Lin, Qiuping Jiang, and Gangyi Jiang. 2017. QoE-guided warping for stereoscopic image retargeting. *IEEE Transactions on Image Processing* 26, 10 (2017), 4790–4805.
[29] Feng Shao, Libo Shen, Qiuping Jiang, Randi Fu, and Gangyi Jiang. 2017. StereoEditor: Controllable stereoscopic display by content retargeting. *Optics Express* 25, 26 (2017), 33202–33215.
[30] Kei Utsugi, Takuma Shibahara, Takafumi Koike, Keita Takahashi, and Takeshi Naemura. 2010. Seam carving for stereo images. In *3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2010*. IEEE, 1–4.
[31] Rafael Grompone Von Gioi, Jérémie Jakubowicz, Jean-Michel Morel, and Gregory Randall. 2012. LSD: A line segment detector. *Image Processing On Line* 2 (2012), 35–55.
[32] Qi Wang and Yuan Yuan. 2014. Learning to resize image. *Neurocomputing* 131 (2014), 357–367.
[33] Qi Wang, Yuan Yuan, and Pingkun Yan. 2013. Visual saliency by selective contrast. *IEEE Transactions on Circuits and Systems for Video Technology* 23, 7 (2013), 1150–1155.
[34] Wenguan Wang, Jianbing Shen, Yizhou Yu, and Kwan-Liu Ma. 2017. Stereoscopic thumbnail creation via efficient stereo saliency detection. *IEEE Transactions on Visualization and Computer Graphics* 23, 8 (2017), 2014–2027.
[35] Kanoksak Wattanachote, Timothy K. Shih, Wen-Lung Chang, and Hon-Hang Chang. 2015. Tamper detection of JPEG image due to seam modifications. *IEEE Transactions on Information Forensics and Security* 10, 12 (2015), 2477–2491.
[36] Weiqing Yan, Chunping Hou, Baoliang Wang, and Laihua Wang. 2017. Content-aware disparity adjustment for different stereo displays. *Multimedia Tools and Applications* 76, 8 (2017), 10465–10479.
[37] Weiqing Yan, Chunping Hou, Yuan Zhou, and Wei Xiang. 2016. Content-and disparity-adaptive stereoscopic image retargeting. *Journal of Modern Optics* 63, 3 (2016), 207–218.
[38] Ting Yin, Gaobo Yang, Leida Li, Dengyong Zhang, and Xingming Sun. 2015. Detecting seam carving based image resizing using local binary patterns. *Computers & Security* 55 (2015), 130–141.
[39] Jin Woo Yoo, Sehoon Yea, and In Kyu Park. 2013. Content-driven retargeting of stereoscopic images. *IEEE Signal Processing Letters* 20, 5 (2013), 519–522.
[40] Dengyong Zhang, Ting Yin, Gaobo Yang, Ming Xia, Leida Li, and Xingming Sun. 2017. Detecting image seam carving with low scaling ratio using multi-scale spatial and spectral entropies. *Journal of Visual Communication and Image Representation* 48 (2017), 281–291.

Received July 2018; revised March 2019; accepted March 2019