Stereoscopic 3D Image Retargeting Quality Assessment

Dandan Li
College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin 150040, Heilongjiang, China

Jun Cao *
College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin 150040, Heilongjiang, China
*Corresponding author(E-mail: 19410381@qq.com)

Yi Liu
Beijing Aerospace Control Center, Beijing 100094, China

Abstract
Image quality assessment is one of the most important ways to accelerate advances of image processing techniques. Stereoscopic 3D image retargeting technique aims to improve the display quality and stereo perception quality during changing the aspect ratio and adjusting stereoscopic depth. However, there is still no objective metric to assess the visual quality of retargeted 3D image. In this paper, we propose a learning-based objective method to evaluate the 3D image retargeting quality. We build the first stereoscopic image retargeting database based on the public datasets and the representative 3D retargeting methods. We conduct subjective test to rank the perception quality for the images in our database. After analyzing the factors that influence the perception quality, we extract new features of quality assessment and fuse them to predict the overall quality score for human’s stereo visual perception. Experiments confirm the effectiveness of the proposed method. The results demonstrate the good consistency between the subjective ranking and the objective assessment scores.

Key words: Quality assessment, Stereoscopic image retargeting, Machine learning.

1. Introduction
In recent years, stereoscopic 3D media becomes increasingly popular because of the good visual effect. There has been significant progress in 3D media content production and 3D display technologies. A variety of stereoscopic display devices are invented to meet people’s need of viewing 3D media. With the development of mobile internet and social network technologies, a stereoscopic image could be shared quickly and displayed on various screens in many scenarios. But different screens have different sizes or aspect ratios. In order to make the most of whole screen and improve the visual experience, stereoscopic image retargeting technique has been proposed and become more and more important. Stereoscopic image retargeting is a technique that addresses human’s visual demand to display 3D images on the devices with different sizes. Traditional retargeting techniques include letterbox (or pillarbox), cropping, scaling etc. They could not completely meet people’s requirement because of their drawback of the missing of some parts of the contents (e.g. fixed-window cropping) or the visual artifacts introduced by over-squeezing the contents (e.g. homogenous scaling). Thus, many content-aware stereoscopic image retargeting techniques were proposed these years, such as [1], etc. By these methods, users could manipulate the depth perception or preserve the important information while adjusting the aspect ratios of input images.

Most previous works of the stereoscopic image retargeting have demonstrated novelty in problem formulation and algorithmic design. However, the performance evaluation of different retargeting methods is generally improvised as most of them resorted to simple visual comparison with a small set of images or conducted small-scale user studies to support their evaluation. Therefore, the necessity to define standardized methods to evaluate the perceived quality of retargeted stereoscopic images is evident. The quality assessment of retargeted stereoscopic images could be not only beneficial to perform the comparison of the retargeting methods but to guide the retargeting process and promote the improvement of retargeting techniques. Image quality assessment is achievable either through subjective testing or through objective metrics. Ideally assessment by human beings with normal ability to view 3D contents is suitable for stereoscopic image retargeting evaluation. But subjective assessment is a time-consuming and expensive approach. Furthermore, the analysis of the obtained results is not straightforward. Therefore, objective methods providing reliable predictions of stereoscopic image retargeting quality are desired. In this paper, the aim of our research is to propose an objective quality assessment method for 3D image retargeting. In the literatures, to the best of our knowledge, much less has been done to assess stereoscopic image retargeting quality objectively. The most relevant works in the image quality assessment area are the objective evaluation techniques for 2D image retargeting quality [2] and stereoscopic image quality [3], which are detailed in Section II. But obviously, both of the two kinds of methods can’t be used to assess the 3D quality.
image retargeting quality. Because a single 2D image can’t generate the disparity, the 2D image retargeting quality assessment methods do not need the ability to evaluate the quality of the perceived depth. On the other hand, the 3D image quality assessment methods are generally classified into three categories: full reference (FR), reduced reference (RR) and no reference (NR). FR methods require all information of original 3D image. RR methods only require partial information. But neither FR methods nor RR methods can be used to evaluate 3D image retargeting quality because they require that test image and its reference image have the same size while retargeting methods aim to change the image size. NR methods evaluate the quality of 3D images without any reference information. But discarding all information of original 3D image, NR methods can’t compute the quality loss caused by the changes of image contents and perceived quality. Therefore, the above methods can’t be used to assess the quality of retargeted stereoscopic images directly, because they are not designed to solve this problem at first and can only take count of parts of factors influencing the 3D image retargeting quality. The performance of some representative methods is tested quantitatively in Section V.

In this paper, we propose a novel objective method to assess the quality of retargeted stereoscopic images. Our goal is to approximate human’s quality assessment of retargeted 3D images. We first build a stereoscopic image retargeting database with subjective testing results. In the experiment, we retarget the original 3D images by different 3D image retargeting methods and take them as the stimulus. The perception quality of each image pair is rated by human viewers on a predefined scale. Thus, we could obtain the subjective quality score for every retargeted 3D image. Given the database, we could analyse the factors that could influence the viewers’ perceived quality. We then extract five features and fuse them to objectively evaluate the overall quality of retargeted stereoscopic images. Fig.1 shows the framework of our proposed method. The extracted features are divided into two types: stereo perception quality and image retargeting quality. Finally, the experimental results show the good consistency between the objective assessments and the subjective rankings. That is to say, the proposed metric can be used to measure human’s subjective perception quality.

Fig. 1. Framework of the proposed quality assessment method.

2. Related Work

In this section, we survey literature on 3D image retargeting techniques, 2D image retargeting quality assessment methods and 3D image quality assessment methods.

2.1. 3D Image Retargeting

The traditional retargeting methods for stereoscopic images include fixed-window cropping, uniform scaling and so on, which is performed without any knowledge about the image contents. Recently many content-aware stereoscopic image retargeting techniques are proposed, which could not only protect the important contents in an image from impairment but also adjust the stereo perception in a certain range.

Image retargeting algorithms can be classified as the discrete method and the continuous method. The discrete methods remove or insert pixels to some region in the images, such as seamcarving or cropping-based methods. Chamaret improve the traditional cropping and scaling method to converse 3D image’s aspect ratio according to the important regions. Basha et al.[2] resize stereoscopic images by using “seam carving” technique and tries to reduce the depth distortion. They take into account the visibility relations between pixels in the image pair and guarantee that the retargeted 3D image is geometrically consistent. Ma et al. [4] proposed visual attention
model and minimize both visual content distortion and depth perception distortion by seam carving method. Lei et al.[3] simultaneously consider seam selecting and seam matching. They preserve the depth based on pixel fusion when retargeting the stereo image pair.

The continuous methods optimize a mapping from original image to its retargeted version in order to constrain the image deformation and improve 3D perception. The warping-based retargeting methods are widely adopted to process 3D images[7]. Fang et al.[8] propose a warping-based method by solving a least-squares energy minimization problem. They perform image retargeting and depth adaptation simultaneously. Ryu et al.[9] decompose the images into layer and warps the contents at the layer-level to guide the image resizing. Lin et al.[1] utilize the information of matched objects rather than that of matched pixels in warping. They consistently preserve both the disparities and shapes of visually salient objects.

In this paper, we will analyse the 3D image retargeting techniques in the literature and introduce as many distortion types as possible into the database.

2.2. 2D Image Retargeting Quality Assessment

Image quality assessment has been an active research area. The quality assessment of retargeted 2D image takes an important part. Fang et al.[8] propose the image retargeting structural similarity map (IR-SSIM) algorithm, which can improve the quality prediction of retargeted images. They also demonstrate it by embedding it into a multi-operator image retargeting process. Ma et al.[4] evaluate the retargeting quality by pairwise ranking method. They resort to learning a ranking model and training the mapping function to give the perceptual quality indexes for the images in the test dataset.

2.3. 3D Image Quality Assessment

Assessing the quality of stereoscopic images is also a challenging issue. Some researchers focus on the influence caused by some kinds of impairments such as blurring, noise, compression etc. But retargeting methods can not cause these impairments. Some researchers consider to measure geometric distortions and visual discomfort in 3D domain, which often occur in 3D retargeting process and influence the final stereo visual quality. We mainly survey the literature on this research area.

Yong et al.[9] improves stereoscopic viewing experience in stereoscopic displays under guidance of their proposed objective assessment metric of visual comfort. Qi et al.[5] evaluate the quality of 3D image by measuring the difference of the original and distorted images' binocular perceptual information, including the distribution statistics of visual primitives in left and right views. Shao et al.[6] propose a learning-based no-reference 3D image quality assessment method. They learn the features by using joint sparse representation and train the quality prediction model to compute the objective quality scores based on feature distribution weights.

3. Database

To our knowledge, few public databases of retargeted stereoscopic images were built in the literature. In order to build the database, we survey the previous works of 3D retargeting techniques. Whether for 2D image or 3D image, when the ratio aspect changes, researchers conduct retargeting process by disposing of the contents of the images. But comparing to 2D image, stereoscopic 3D image brings more stereo visual experience to viewers. Retargeting for 3D image could not only introduce artificiality but also impair the stereo perception. Therefore, we process the original 3D images by using different retargeting methods, which can generate new retargeted images containing different types of distortion and stimulate viewer’s vision more fully in our subjective test. On the other hand, for image retargeting techniques, one of the challenges is posed for some types of image contents containing dense information or structures, such as face, body, building, etc. For the images with different contents, the same retargeting method could produce the outputs bringing the different visual experiences. We collect the source images from the retargeting papers and 3D image datasets, which include various attributes (people, animal, natural scenery, etc. Seeing Fig.2). Thus, we process these images by using various retargeting techniques or obtain the proposed public results of retargeting papers. Based on these outputs, we build the retargeted stereoscopic image database finally.

3.1. Retargeting Methods

As mentioned in Section.II, the content-aware retargeting methods are classified as discrete or the continuous. In our study, we select warping-based methods and seam-carving-based methods to represent the two types of retargeting techniques respectively. When retargeting the 3D image, the distortions mainly result from three problems: saliency detection error, disparity computing error and model flaw. First, as the name suggests, the content-aware retargeting methods must be sensitive to the important areas in images. Researchers often achieve the goal by computing saliency map before resizing step. We adopt three saliency detection methods with different characteristics or edit the saliency map manually to simulate the error. Second, the problem of computing disparity is very hard and unresolved perfectly in the literatures yet. The stereo perception quality of retargeted image cannot be guaranteed completely with the errors of disparity values. In our work, we compute the disparity value
by SURF algorithm. Third, all 3D image retargeting models have their own shortcomings and cannot be applied to every image. In our study, we use the warping-based methods proposed by Lin et al.[1] and the seam-carving-based methods proposed by Basha et al.[2].

![Image of original images](image-url)

**Fig. 2.** The original images in our test.

Besides the content-aware retargeting methods, we also use the traditional retargeting methods: cropping and scaling, in order to introduce more distortion types. We set the cropping windows manually to investigate the perceptual tradeoff between picture deformation and content loss.

Above all, we use four representative retargeting methods to process these source images: cropping(CR), linear scaling(LS), seam-carving-based method(SC) and warping-based method(WP). Sample results we used in our study are shown in Fig.3.

![Sample results of retargeting methods](image-url)

**Fig. 3.** Example of retargeting a stereoscopic 3D image in our database.

3.2. **Stimuli**

We collect 56 stereoscopic 3D images from datasets CMU/VASC and NVIDIA and stereoscopic image repository on Flickr, which contains various attributes including people, animal, cartoon, natural scenery, car, building, etc. (seeing in Fig.2.). With these source images covering so many attributes, the performance of the selected retargeting methods are challenged. It is easy to produce the retargeting results full of various distortions.

Without loss of generality, we set the target aspect ratio as $67\% W \times H$, where $W$ and $H$ indicate the width and height of the original images respectively. Furthermore, to simplify the experiment, we require that the retargeting method should preserve the stereo perception of the original image, namely, the original disparity values. We process the source images by using the above four 3D image retargeting methods. In Fig.3 we show the four retargeted results of a sample image in our database. The main content (bird) in the foreground is partly removed by the cropping method in Fig.3.(g), while the disparity is preserved completely. On the contrary, retargeting by the scaling method, all the image contents are preserved but the disparity values decreased.
uniformly (seeing the anaglyph in Fig.3.(h).). The results processed by WP and SC are shown in Fig.3, in which we find obvious deformation marked by red box in Fig.3.(d), Fig.3.(e), and disparity value change marked by yellow box in Fig.3.(j). Finally, the database contains 224 retargeted stereoscopic images as stimuli, which is available online.

3.3. Subjective Testing

Many recommendations for subjective evaluation of visual stimuli have been issued by the International Telecommunication Union (ITU). But there is no applicable standard or recommendation for 3D image retargeting quality assessment. We reference these methodologies and further customize our subjective test.

![Fig. 4. (a) and (b) are the left view and the anaglyph of two 3D images respectively. Which is the original image, (a) or (b)?](image)

In this study we collect the subjective evaluation results for the retargeted 3D images. Being different to evaluate the subjective image-quality of monoscopic pictures or the stereo perception quality of 3D images, the change of picture or 3D experience sometimes is difficult to perceive. For example, in Fig.4, we hardly to tell which the original image or the retargeted one is when looking at only one of them. Therefore, both the test image and the reference image should be indicated for subjects before they evaluate the quality of the test image. We ask subjects to consider two questions:

1. What do you think about the quality of the test image itself?
2. If your picture (the reference image) is processed and become the test image, please grade the acceptability score.

The aim of the two questions is to guide the subjects to evaluate both the test image’s self-quality and its comparative quality referencing visual perception of the original image. In our experiment, subjects give a composite score for the quality of the retargeted 3D image.

We select the simultaneous double stimulus for continuous evaluation (SDSCE) method. This method is more suitable for 3D image retargeting quality assessment than the DSCQS method. For 3D image, the switching operation in DSCQS method could cause visual fatigue [11], which could cause subject’s bad evaluation. For the same reason, we ask subjects to speak out a score instead of writing it down, which can avoid them accommodate their eyes frequently. We record all of the evaluation scores for subjects.

Twenty persons with normal stereo vision ability took part in the subjective testing. In order to reduce the effect of the viewer fatigue, we divide the 224 images into 4 sessions randomly. Every session contains 56 images.
We conduct the experiment using simultaneous double stimulus for continuous evaluation (SDSCE) method in full screen. The order of the images is random for the subjects. Five-point quality scale is used to evaluate the quality of the retargeted stereoscopic images: 5-"Excellent", 4-"Good", 3-"Fair", 2-"Poor", 1-"Bad". The subjects give a score referencing visual perception of the original 3D images.

We use a 3D monitor (Acer GD245HQ, 1920×1280 pixels, 300cd=m\(^2\), 23.6 inch) and nVIDIA GeForce 3D VISION devices in our experiment. The viewing distance is set to 3H(about 80cm), where H is the picture height. We present two stereoscopic images in the monitor at a time, side by side, with black background in full screen mode. The subjects are free to see the next or previous 3D image pairs until they have formed an opinion. Due to the possibility of eye fatigue, there is one minutes break between the sessions.

Because of the individual differences in visual perception or some unforeseen problems, the subjective assessment results for an image could be inconsistent. But a controversial result cannot be used as a convictive quality score of an image. We employ the quartiles of the subjective scores for every stereoscopic image to check the validation of evaluation (Fig.5). If the interquartile range (IQR) is larger than 2, we consider the image’s subjective quality is controversial and the subjective score is invalid. Finally, we obtain 215 retargeted stereoscopic images. These images can be used to conduct objective testing.

![Fig. 5. Interquartile range map of subjective assessment score in our experiment.](image)

4. The Proposed Approach

In this section, we conduct our objective test based on the database. The question we consider first is what methodology we should follow to model the objective quality assessment for retargeted 3D image. In the literature we find that many research works have been done on the image quality assessment including the above-mentioned 2D image retargeting quality and 3D image quality in Section.II. So we think that it could be easier and more efficient to study the 3D image retargeting quality heuristically by using the existing knowledge.

There are two traditional issues in objective image quality assessment. The first one is the extraction of appropriate features. The good features can reveal the key factors influencing the image quality. In fact, measuring a feature quantitively is usually implicitly evaluating one or more aspects of image quality objectively. Just for this reason, various features need to be integrated in a quality assessment method to cover the quality factors. So the second one is the fusion of the extracted features, which generates a number to represent quality score. The correlation level between the objective score and the subjective ranking could be different because of the adopted fusion method. Therefore, our aim is to extract the good features and integrate them by proper fusion method.

Due to the characteristic of retargeted stereoscopic images, we consider two issues to find the ways to extract features.

1. What is the impact of retargeting methods on viewer’s experience?
2. Independently of retargeting techniques, which factors should be taken into account when evaluating the impairment of perception quality?

We thought about the first question in order to find what modifications of the retargeting method decrease the quality of viewer’s experience. The modified areas in the image implies the effective features we need most, such as the distortion or loss of the image contents and so on. On the other hand, looking at a stereo image, viewer maybe intuitively find some artificial areas or feel uncomfortable before they compare the retargeted image with the original image, which got our attention when we conducted the subjective test. If we extract these kinds of features, we can measure the perceptive quality of the retargeted 3D image quickly and directly.
Based on these considerations, we try to evaluate retargeted stereoscopic images from two aspects: **stereo perception quality** and **image retargeting quality**. Stereo perception quality describes the changes of the stereoscopy’s distribution and intensity. We employ depth similarity and disparity excessiveness to assess the stereo perception quality. We consider that it includes the influences on viewer’s visual experience both from the retargeting processing and from the retargeted 3D image itself. Image retargeting quality represents the changes of the picture’s gross information content. We observe the retargeted results in our database. We find that, in order to meet the new aspect ratio, in the original image, some contents are either given up to be removed, or transformed according to the mapping algorithm. So, we decide to assess the image retargeting quality by exploiting picture completeness, local distortion and global distortion in the retargeted images.

![Example of extracting depth similarity feature.](image)

**Fig. 6.** Example of extracting depth similarity feature.

### 4.1. Stereo perception quality

As a stereoscopic image, 3D perception is its key characteristic. The stereo perception quality is certainly one of important parts of stereoscopic image retargeting quality. The original stereo perception should be modified carefully, which should preserve the original depth structure and make the output’s depth looked comfortable. Two features are extracted to represent the stereo perception quality in our proposed method: depth similarity and disparity excessiveness.

**Depth similarity.**

Depth information in a 3D image is the source of the stereo perception. All of the objects in the image can generate the feeling of relative depth. They are located in the different depth of field because of their different depth value. In an image, the objects in the background always have the larger depth values and look further away from viewer. Conversely, the foreground contents have the smaller depth values and look in front of the background. Depth is helpful for viewer to understand image. Although depth cues include occlusion, size, etc besides the depth value, most of retargeting methods manipulate depth perception mainly by changing depth value rather than other cues. In fact, in an 3D image, binocular disparity is one of the most accurate and efficient depth cues, which works best in the near distance especially.

We consider the basic depth constraints in the retargeting methods in order to find a way to measure the changes of depth structure. With the different depth, the whole image can be divided into some layers in depth. The 2D or 3D image can be decomposed into layers according to the depth information. The basic constraint of stereoscopic image retargeting techniques is to preserve the relative depth or layers of image contents. Therefore, we propose a new assessment feature called “Depth Similarity” \( S_D \) to measure the depth preservation before and after retargeting. The depth similarity describes the inheritance of binocular disparity distribution during the retargeting processing.

Given the original stereo image pair \( \{ I_L, I_R \} \) and its retargeted results \( \{ I'_L, I'_R \} \), we compute the normalized disparity maps \( (D, D') \) respectively. Using SURF algorithm, we match the key points in the original image and the retargeted image, which can help us to find the corresponding positions between them. Then we will check the depth changes at these corresponding key points by comparing the depth values. To enhance robustness, we
use the disparity values of the patches where the points are. In our experiments, patch size is set as 11 × 11. We compute the average change of disparity in the corresponding patches. Considering the human vision system is more sensitive to salient regions, not all patch pairs have the same importance level. In some insignificant areas that people do not notice, they certainly will not perceive the depth there. Thus, to improve the proposed depth similarity feature, we compute the saliency of the original image and weight the depth changes using average saliency values of patches (si).

\[
S_D = \frac{\sum_{i,j \in S_{i,j}} |D_i - D_j|}{\sum_{i,j \in S_{i,j}}}
\]

where \(N\) is the number of patch pairs. \(P_i\) is one of patches in the original image, whose corresponding patch in the retargeted image is \(Q_i\). \(D_i\) and \(D'_i\) is the depth values of \(P_i\) and \(Q_i\) respectively. \(s_i\) is the average saliency value of the pixels in \(P_i\). When a patch’s depth value is modified after retargeting, if the patch is more important or its depth value changes greater, it has more influence on the image’s depth similarity.

Fig.6 shows an example of extracting depth similarity feature. Fig.6.(a) is the left view of the original 3D image. Fig.6.(c) is the left view of the retargeted 3D image. Fig.6.(b) and Fig.6.(c) show the patch correspondance. Fig.6.(d) presents the original depth values of the patches. And the depth values after retargeting are shown in Fig.6.(e). Fig.6.(f) shows the saliency of each patch.

**Disparity excessiveness.**

Stereopsis is the perception of depth that is constructed based on the binocular disparity. Because of the binocular suppression, human visual system (HVS) can tolerate some level of the accommodation-vergence conflict caused by the horizontal disparity and perceive the comfortable stereo vision. However, when the conflict exceeds some tolerance level, HVS can perceive the loss of accommodation and/or loss of binocular fusion. Various discomfort symptoms could be induced.

For retargeted stereoscopic images, the relative depth could be preserved by control the depth similarity. But it is not enough to assure the comfortable stereo perception. The absolute depth value should be constrained as well. The excessive absolute disparity still colud be generated by some retargeting methods, such as uniform stretching, warping, etc. Therefore, we extract the disparity excessiveness feature to evaluate 3D image quality.

It is unsuitable to compute the binocular disparity by using the pixel number of image because that the display devices are different. The visual discomfort is related with the retinal disparity rather than the pixel disparity [10]. The threshold of retinal disparity is 30 arcmin. It could cause the accommodation-vergence conflict if the disparity exceeds the threshold in a 3D image. Thus, given the viewing distance and the screen’s pixel density, we can compute the threshold of the horizontal pixel disparity \([D_L;D_U]\). (The inter-ocular distance is set as 65mm generally). Furthermore, the excessive disparity at the saliency point could get viewer’s more attention and cause the discomfort more easily. We can express the horizontal disparity feature \(S_H\) by measuring the importance of the pixels with excessive disparity.

\[
S_H = 1 - \frac{\sum_{i=1}^{N_i} S(i)}{\sum_{j=1}^{N_i} S(j)}
\]

\(N_i\) is the number of the pixels whose disparity values exceed the threshold. \(N_i\) is the number of all pixels in the image. \(S(i)\) is the saliency value of each pixel.

4.2 Image retargeting quality

In order to meet the need of changing aspect ratios, stereoscopic image retargeting techniques often cause the distortion of the image contents. These kinds of distortion are different from the artifacts induced in the coding or transmission step, such as adding noise, blurring and compression, etc. They are usually generated because of moving, copying or removing pixels during retargeting processing. There are three main distortion types of image retargeting. The first one is the loss of image content. When retargeting by removing some parts of the image, a good retargeting method should preserve as much important contents in the original image as possible. The retargeting methods that most likely cause loss of image information is cropping-based method. When the important regions are large enough, cropping will inevitably discard some image contents.

The second distortion type is local region distortion. When researchers try to preserve all of the information in the original image, the shape of the unimportant areas will be sacrificed in the retargeting step to protect the important contents from transformation. This kind of distortion often appears in the content-aware retargeting methods, especially the continuous type methods. But because of some reasons, such as retargeting algorithm robustness, saliency detection error and so on, the distortion could get viewer’s attention and cause to feel disgust, which reduces the image quality directly.

286
The third distortion type is global structural distortion, which leads to unpleasing shape deformation throughout the whole image. Sometimes this kind of distortion is not easy to be perceived because of the objects in the image all are lack of texture such as clear sky or white wall, or are full of texture such as fabric or patterns, etc. But retargeting this type of images is of no challenge in fact. Most of the images in our lives are easy to be influenced by the global structural distortion. It occurs when an image is over-squeezed or overstretched after retargeting.

Based on the above analysis, we measure these kinds of distortion by improving the existing algorithm and propose three features: **Picture Completeness, Local Distortion, Global Distortion**.

We use bidirectional similarity (BDS) algorithm to measure picture completeness feature and local distortion feature. The level that the retargeted image represents the visual data of the original image is defined as picture completeness. If all of the information in the original image are found in the retargeted image, we can conclude that the picture completeness is very good. In our subjective test, we observed that viewer do not hope to lose any parts of the original information in fact. Second, the level that the retargeted image introduce new visual local artifacts that were not observed in the original image is measured as local distortion. We consider that the new introduced artifacts in the retargeted image actually are the excessive transformation of the contents in the original image. We define them as local distortion.

However, BDS algorithm might make mistakes in the background regions of images because of the global patch comparison. There are two reasons for us to reduce the quality weight of the background part. The first one is that the background is usually not an important region for viewer. The other is that in many cases the texture of the background region is so simple that the BDS distance is very small there. Additionally the background occupies a large parts of image, which could reduce the mean BDS distance of the whole image and hide the bad quality in some key area. It will invalidate the quality measurement. Furthermore, in the retargeted image, not all local distortions are important enough to grab attention. The unattractive local distortion will overrate the bad impact on the overall image quality. Therefore, we employ the saliency value of images and set a threshold to filter the insensitive distortion regions for viewers. The picture completeness feature $S_P$ measures the loss of important regions in the two views of the original stereoscopic image. For an example, see Fig.8.

![Fig. 7. Framework of our proposed method.](image-url)
Fig. 8. Example of picture completeness feature and local distortion feature.

\[ S_p = \left( \frac{\sum_{i=1}^{N_{r1}} B_i}{N_{r1}} + \frac{\sum_{j=1}^{N_{r2}} B_j}{N_{r2}} \right) / 2 \]

where \( N_{r1} \) is the number of regions whose saliency value is larger than the threshold in the left view, and \( N_{r2} \) is that in the right view. \( B \) is the completeness level of a region. In our experiments, the saliency threshold \( s_t \) is 0.5 \((s_t \in [0, 1])\).

The local distortion feature \( S_L \) measures the distortion level of the important regions in the views of the retargeted stereoscopic image. (See Fig.8.)

\[ S_L = \left( \frac{\sum_{i=1}^{N'_{r1}} B'_i}{N'_{r1}} + \frac{\sum_{j=1}^{N'_{r2}} B'_j}{N'_{r2}} \right) / 2 \]

where \( N'_{r1} \) is the number of regions whose saliency value is larger than the threshold in the left view, and \( N'_{r2} \) is that in the right view. \( B' \) is the coherence level of a region.

In Fig.8, (a) is the left view of the original 3D image. (b) is the saliency map of (a). (c) and (d) are the retargeted results of different methods. (e) presents (c)'s picture completeness w.r.t. (a). (f) presents (d)'s local distortion level w.r.t. (a).

The algorithm termed “Spatial Envelope” is described as a set of perceptual properties (naturalness, openness, roughness, etc.). The spatial envelope properties provide a global description of the scene. We adopt it to measure the global structural distortion for retargeted stereoscopic images. Thus our global distortion feature \( S_G \) is defined as

\[ S_G = 1 - (G_L + G_R) / 2 \]
where $G_L$ and $G_R$ are the global distortion level of left view and right view respectively.

### 4.3. Fusion of features

Besides the feature extraction in Section 3, another key issue of objective image quality assessment is to find an appropriate model to fuse the features and closely predict subjective perceived quality. In this work, we formulate the retargeted stereoscopic image quality assessment as a regression problem based on the proposed features.

Because of the complexities of human stereo vision and psychology, we cannot infer the relationship between the proposed features and the subjective perception. Therefore, we attempt to accomplish the quality assessment using machine learning method. In this paper, we choose radial basis function (RBF) neural network to learn the pattern of retargeted stereoscopic image quality evaluation. RBF network has three layers: an input layer, a hidden layer with a non-linear RBF activation function and a output layer. An RBF neural network with $K$ hidden layer neurons can be described by

$$y = \sum_{k=1}^{K} \omega_k \theta_k(x) \tag{6}$$

where $x$ is the input vector, and $y$ is a scalar function of $x$. $K$ is the number of hidden neuron. $\theta_k$ is the output value of the $k$-th hidden neuron, whose weight for the output layer is $\omega_k$. By choosing optimal weights, the output result is optimized. An RBF network with enough hidden neurons can approximate any continuous function with arbitrary precision.

We divided the 215 retargeted stereoscopic images into two parts: 85% of them (183 images) are used as training set and 15% of them (32 images) are used as testing set. In the training process, we train the RBF neuron network using the features of the images in the training set and their subjective quality scores. After learning, we obtain the assessment model. According to the properties of RBF neural network, the model we obtained is a fusion of these features. In the testing process, the RBF neural network takes the features of the images in testing set as inputs and gives a score for each image by using the model we have learned. The output scores represent the objective quality of these retargeted stereoscopic images.

### 5. Experimental Results and Analyses

Since we have obtained both subjective and objective quality rankings of the retargeted stereoscopic images in the testing set, in order to verify the proposed model’s performance and effectiveness, we use the following four indexes to measure the level of agreement. The first index is the Pearson linear correlation coefficient (PLCC). It provides an evaluation of prediction accuracy. The second index is the Spearman rank order correlation coefficient (SROCC). It is considered as a measure of prediction monotonicity. The third index is the Kendall rank correlation coefficient (KRCC), which evaluate the consistency of prediction results. The fourth index is the root-mean-squared error (RMSE). For an ideal match between the objective prediction scores and the subjective quality scores, PLCC=1, SROCC=1, KRCC=1 and RMSE=0.

<table>
<thead>
<tr>
<th></th>
<th>SROCC</th>
<th>PLCC</th>
<th>KRCC</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMD</td>
<td>0.0698</td>
<td>0.0470</td>
<td>0.0450</td>
<td>0.3217</td>
</tr>
<tr>
<td>SIFT-Flow</td>
<td>-0.0300</td>
<td>0.0076</td>
<td>-0.0216</td>
<td>0.2511</td>
</tr>
<tr>
<td>BDS</td>
<td>-0.0425</td>
<td>-0.0662</td>
<td>-0.0286</td>
<td>0.2705</td>
</tr>
<tr>
<td>DE</td>
<td>0.0527</td>
<td>0.1123</td>
<td>0.0388</td>
<td>0.3280</td>
</tr>
<tr>
<td>DS</td>
<td>0.1732</td>
<td>0.1739</td>
<td>0.1201</td>
<td>0.2841</td>
</tr>
<tr>
<td>CSVT14</td>
<td>0.0498</td>
<td>0.0884</td>
<td>0.1201</td>
<td>0.0341</td>
</tr>
<tr>
<td>Ours</td>
<td>0.8194</td>
<td>0.6856</td>
<td>0.6757</td>
<td>0.1581</td>
</tr>
</tbody>
</table>
As mentioned in Section I, we analysed why the objective evaluation techniques for 2D image retargeting quality and stereoscopic image quality cannot be used to assess the 3D image retargeting quality. To demonstrate this inference, we compute the evaluation scores by using those techniques and compare the results with the subjective scores. The values of the above indexes can be used to verify their assessment performance. The 2D image retargeting quality assessment methods that we select are: Earth Mover Distance (EMD), SIFT-Flow, Bidirectional Similarity (BDS), which are representative metrics in the literature. But, to our knowledge, no public accessible no-reference assessment technique is provided to evaluate the quality of stereoscopic images. Therefore, we code Ryu et al.’s algorithm (CSVT14) which is a NR assessment method with high performance for the stereoscopic image [9]. The objective score generated by this algorithm can be used to measure the stereo visual perception quality of 3D image. Besides, for comparison, we also use our proposed two stereo perception quality features to evaluate the quality of stereoscopic images: Depth Similarity (DS) and Disparity Excessiveness (DE). The results are shown in Table I and Fig.9.

From these experiment results, we can observe that our method outperforms other objective assessment metrics significantly. The traditional quality assessment techniques for 2D image retargeting could hardly predict the quality of retargeted 3D images. It shows a very weak positive correlation between these metrics and the subjective quality. On the other hand, by using the metrics for stereo perception quality, the prediction results are more encouraging. It is evident that the stereo perception quality metrics better agree with the user labeling in comparison to the 2D image retargeting quality metrics, although the correlation values are not high. The algorithm of CSVT14 achieves poor performance. We infer that the reason is CSVT14 algorithm may not apply to measure the quality impairment caused by disparity distortion, but be good at predicting perpetual quality influenced by blurriness and blockiness.

Through the above analysis, we consider whether the stereo perception quality is more relevant to the final quality of retargeted 3D image in the experiment. We compute the correlation values between the subjective grading and the objective scores obtained by using our five features respectively. The results are shown in Fig.10. According to the correlation indexes, almost all of the five features are positively correlated with human labeling. But the depth similarity feature’s correlation value is evidently more consistent with our final prediction. Therefore, up to a certain extent, the results show that subjects attach more importance to the change of the perceived depth during the 3D image retargeting processing.

Finally, our features of image retargeting quality and stereo perception quality all contribute to the objective grading. Furthermore, by combining these features in the experiment, our model does capture the distortions and quantify their impact on visual perception. It achieves better performance to measure human’s visual perception quality of 3D image retargeting.

6. Conclusions

We present the first study on stereoscopic 3D image retargeting quality assessment. In this paper, we collect 3D images from the public datasets and online website, which covering many attributes. We employ four different representative retargeting methods to produce the retargeted results and introduce as many distortion types as possible. Based on these outputs, we build the first retargeted stereoscopic image database. With the image database, we conduct the subjective test and consider the factors that influence the quality of retargeted 3D image. In order to represent the factors, we analyse the changes of 3D image during the retargeting processing and extract five features of quality assessment for retargeted 3D image: depth similarity, disparity excessiveness, picture completeness, local distortion and global distortion. We combined these features and propose a new quality...
assessment method for retargeted stereoscopic images using radial basis function neural network, which is an attempt in stereoscopic image retargeting case. The predicted score describes the overall perceived quality of a retargeted 3D image. It is demonstrated by experimental results that the proposed method achieves a good performance and outperforms other objective quality assessment metrics.

Although the metric presented in this paper shows satisfied performance, there is still room for improvement. Given more training data and more subjective evaluation results, the stereoscopic image retargeting database will become larger and contain more complete distortion types, which can help researchers to extract more effective features and generate the models that can better simulate human 3D visual perception. Many works could be done in the future.

Acknowledgements
This work was supported by the Program of National Natural Science Foundation of China (CN) [Grant No. 31700643, 51801023], the Fundamental Research Funds for the Central Universities [Grant No. 2572016EBC3].

References