Signal Processing 150 (2018) 204–214

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Blind stereoscopic 3D image quality assessment via analysis of naturalness, structure, and binocular asymmetry

Guanghui Yue [a,c], Chunping Hou [a], Qiuping Jiang [b,c,*], Yang Yang [a]

[a] School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China
[b] School of Information Science and Engineering, Ningbo University, Ningbo 315211, China
[c] School of Computer Science and Engineering, Nanyang Technological University, 639798, Singapore

ARTICLE INFO

Article history:
Received 30 November 2017
Revised 4 April 2018
Accepted 19 April 2018
Available online 21 April 2018

Keywords:
Stereoscopic 3D image
Quality assessment
No reference
Binocular vision
Natural scene statistic (NSS)
Asymmetric distortion

ABSTRACT

Over recent years, stereoscopic three dimensional (S3D) images have grown explosively and received increasing attention. Quality assessment, as the fundamental problem, plays an important role in promoting the prevalence of S3D images as well as the associated products. In this paper, an effective blind quality assessment method of S3D images is proposed via analysis of naturalness, structure, and binocular asymmetry. To be specific, given that natural images obey certain regular statistical properties, natural scene statistic (NSS) features of left and right views are first extracted to quantify the naturalness. Second, by considering binocular visual characteristics, statistical features are extracted from a created cyclopean map. Moreover, gray level co-occurrence matrix (GLCM) is utilized to capture quality-sensitive features from the cyclopean phase map. Third, to quantify the asymmetric distortion, a simple but effective measurement is utilized, i.e., calculating the similarity between left and right views as well as statistical features of their difference map. Finally, all extracted quality-sensitive features are combined, and trained together with the subjective ratings to form a regression model using support vector regression (SVR). Experimental results on four publicly available databases (two symmetrically distorted databases and two asymmetrically distorted databases) demonstrate that the proposed method is superior to several mainstream image quality assessment (IQA) metrics.

1. Introduction

Owing to the rapid development of multimedia and networking technologies in the past decades, the volume of digital image data has been growing explosively and played increasingly important role in daily life. However, images inevitably suffer from various distortions during acquisition, transmission, processing, and storage stages, causing unpleasant visual quality of experience. Image quality assessment (IQA) metrics, developed to evaluate or monitor image quality, show great potential on controlling and improving the execution of image processing systems, such as compression, enhancement, and transmission. To date, quantities of mature IQA metrics for two-dimensional (2D) images have been reported [1–8]. Recently, stereoscopic three-dimensional (S3D) multimedia services become popular and greatly increase the immersive experience by enabling depth perception compared to traditional 2D visual experience [9]. Undoubtedly, quality assessment of S3D images becomes more complex as more perceptual issues should be considered: visual fatigue/discomfort induced by binocular depth information [10–12] and annoying experience caused by distortions [13,14]. As the interaction between the visual fatigue/discomfort and the distortion is complex and it has not been fully understood at the current stage. This paper mainly solves the quality assessment problem of S3D images from the viewpoint of image distortions as previous works did [13–16].

Broadly speaking, quality assessment metrics can be divided into two categories: subjective method and objective method. The former case relying on subjective evaluation is more effective, reliable and is usually treated as the ground truth of the latter case. However, it is time-consuming, expensive and unsuitable for online implementation [17]. By contrast, the latter case overcomes these shortcomings and has been widely investigated. Generally, objective methods can be further categorized into full reference (FR), reduced reference (RR), and no reference (NR, also called as blind reference) according to the participation of the reference information [1]. To date, many works have been reported with rather com-

* Corresponding author at: School of Computer Science and Engineering, Nanyang Technological University, 639798, Singapore.
E-mail address: n1609861g@e.ntu.edu.sg (Q. Jiang).

https://doi.org/10.1016/j.sigpro.2018.04.019
0165-1684/

petitive performance on evaluating 2D image quality. Among them, the FR metric, assuming that the reference and distorted images' information is completely known, is the most mature and prevalent ones [2,18–20]. To deal with the problem where only part of the information is sufficient for IQA, RR metrics was proposed and get rapid development [21]. However, due to their intrinsic dependence on reference information, both FR and RR metrics can only be used in limited situations and are powerless to deal with situations, where the reference information cannot be obtained. Therefore, how to design an effective reference-free metric (i.e., NR metric) has received incremental attentions from researchers and gradually become the mainstream direction [22–27].

Different from traditional 2D images, quality assessment related to S3D images is more challenging and various factors, such as 2D image quality, binocular asymmetry, depth information, etc., should also be considered together. To cope with this challenge, many works have been conducted objectively through FR and RR measures. One broadly adopted strategy is to apply existing 2D IQA metrics on both left and right views individually, and then integrate them together to infer the overall quality. Unfortunately, such simple operation usually produces moderate performance. Actually, the perceptual quality of S3D images mainly depends on the merged effects of binocular vision, such as binocular fusion, binocular suppression, and binocular rivalry. Therefore, effectively simulating the binocular visual characteristics may be particularly beneficial to design S3D IQA models. For example, Bensalma et al. proposed a FR S3D IQA metric by calculating the binocular energy difference between the reference and corrupted images [16]. Specifically, the binocular energy was calculated by simulating the properties of simple and complex cells in the visual cortex. Considering the binocular fusion effect, Chen et al. [14] first synthesized a cyclopean view from the stereopair and its associated disparity map. Then, the S3D IQA model was built by applying a traditional 2D FR IQA metric on the generated cyclopean view. According to physiological discoveries on binocular vision, Lin and Wu [15] incorporated binocular integration behaviors (i.e., binocular combination and binocular frequency integration) into existing 2D models to enhance the ability in evaluating S3D images. In [28], Wang et al. developed a 2D-to-3D pooling scheme and designed a binocular rivalry inspired multi-scale model to predict the quality of S3D images. Shao et al. [29] proposed a dictionary based FR metric by considering binocular receipted field properties. Besides, they also constructed a phase-tuned visual codebook and a phase-tuned quality lookup from the binocular energy responses [30]. In [30], the image quality was estimated by averaging the largest value of all patches' qualities, which were obtained via searching the constructed codebook and lookup. Assuming that visual primitives can effectively represent image information, Qi et al. [31] proposed a RR metric. To be specific, they first calculated the entropies of the left and right views (which were obtained based on the estimated coefficients of visual primitives via sparse coding) to represent monocular cue. Then, the mutual information of left and right views was calculated to represent binocular cue. Finally, the cues' differences between the reference and distorted images were fused by a regression function to build the final quality assessment model. Likewise, Wan et al. [32] utilized the same framework as [31] and reported a RR metric by dividing the visual primitives into three categories, i.e., DCprimary, sketch and texture.

Unlike those reference-based metrics mentioned above, NR IQA metrics for S3D images (without reference information in hand) are more difficult to design. Actually, NR IQA metrics have broader application scope compared to the reference-based ones. Although challenging, some progresses have also been achieved in this area [33]. For instance, Liu et al.[34] proposed a S3D IQA model by fully considering the impact of binocular fusion, rivalry, suppression, and a reverse saliency effect on the perception of distortion. By simulating binocular visual system, Zhou and Yu [35] extracted the local binary pattern (LBP) on binocular rivalry response and binocular energy response maps to infer image quality. Besides, they also simulated the simple and complex cells' response maps based on which the LBP descriptors were extracted as the quality-sensitive features [36]. As a stereopair has two views (i.e., left and right views), it inevitably suffers from asymmetric distortions, hereby posing a severer challenge [37]. However, the aforementioned methods ignored the measurement of distortion asymmetry. To cope with this challenge, a few works have been reported. Jiang et al. [33] addressed the quality assessment issue of S3D images using a three-column nonnegativity constrained sparse autoencoder framework, which considered both monocular view quality and the interaction of two monocular views. Shen et al. [38] designed a NR metric for asymmetrically distorted stereoscopic images by utilizing combined model. Specifically, they used the difference map between the Gabor responses of left and right views to reflect the asymmetric information. Shao et al. [39] first judged whether the degree of distortion asymmetry between the left and right views of a stereopair was subjectively noticeable, and then integrated this indicator as a guidance into a supervised dictionary framework for quality evaluation of stereopairs.

Overall, the success of the above reviewed works benefits from the simulation of binocular visual characteristics (e.g., binocular fusion, binocular rivalry and binocular suppression) and the extraction of quality-sensitive features (e.g., LBP, entropy, energy, etc.). However, naturalness, which has been proved to be closely related to image quality, is almost ignored by the aforementioned methods. Although some works applied naturalness as a quality measurement, they did not consider binocular visual characteristics [40,41]. Besides, only part of these works made specific efforts to handle the challenging binocular asymmetric distortions. Actually, evaluating asymmetric distortions is more meaningful in many practical S3D applications. In this paper, an effective blind S3D IQA method is proposed by comprehensively considering properties of binocular vision, image naturalness, and distortion asymmetry. The contributions of this work are three-fold: (1) Towards simulating binocular combination properties, two individual monocular views are synthesized to form a cyclopean image whose mean subtraction and contrast divisive normalization (MSCN) coefficients in spatial domain are fitted using general Gaussian distribution (GGD), then the associated shape and scale parameters are extracted to quantify image naturalness. Compared to competing metrics, the proposed method considers both binocular visual characteristics and image naturalness together. (2) Since phase information reflecting the critical structures in an image is more sensitive to distortion, the quality-sensitive features are captured on the phase component of the cyclopean view via gray level covariance matrix (GLCM). This is a pioneering attempt to solve S3D IQA problem by extracting GLCM on the phase map of cyclopean view. (3) Towards portraying the unique role of binocular asymmetry in affecting the quality of S3D images, the proposed method proposes to compute the similarity between gradient maps of left and right views as well as GGD parameters of their difference map to estimate the degree of distortion asymmetry. Experimental results on public 3D image databases demonstrate that the proposed method is superior to existing mainstream FR and NR IQA metrics.

The rest of the paper is organized as follows. Section 2 gives a detailed description about the proposed blind S3D IQA method. Section 3 presents the experimental results on four publicly available S3D databases and corresponding analyses. Finally, Section 4 concludes the work and gives future directions.

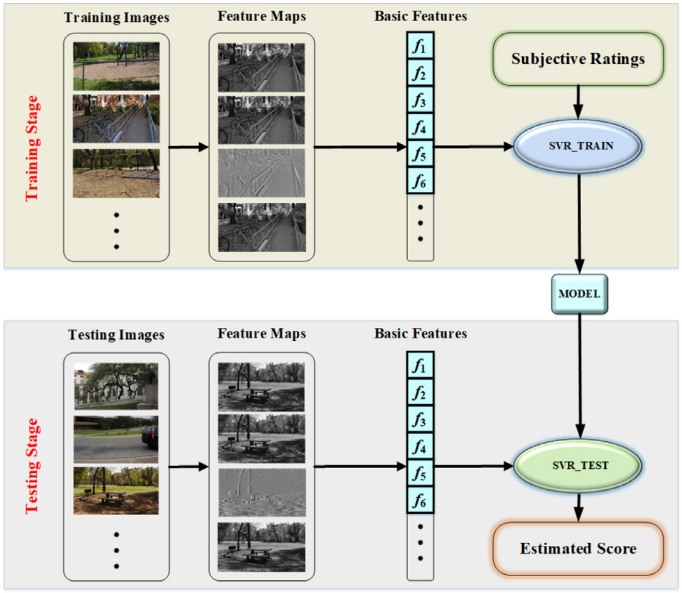

Fig. 1. Framework of the proposed method. For the sake of simplicity, only the left view of stereopair is shown.

2. The proposed blind S3D IQA method

Fig. 1 depicts the framework of the proposed blind S3D IQA method, which consists of training stage and testing stage. In training stage, quality-sensitive features are extracted primarily. Then, a regression model is learned to build the relationship between these features and their subjective ratings (e.g., mean opinion score (MOS) or different mean opinion score (DMOS)) via SVR. In testing stage, the quality-sensitive features are first extracted for test images. Then, image quality scores of these images are estimated by feeding the extracted features into previously trained regression model. In this paper, the quality-sensitive features are composed of NSS features extracted from the stereopair and its cyclopean view, GLCM-based features extracted from the cyclopean phase map and asymmetry-related features. In what follows, each feature component are described in detail.

2.1. Natural scene statistics

Although distinctive in contents, natural images inherently obey a particular statistical characteristic [42], which is measurably modified by the presence of distortions. Motivated by this, previous works have utilized and analyzed NSS in the transform domain and spatial domain [22,23]. In this study, NSS is only considered in spatial domain for the sake of computational efficiency. Given an image I with the size of $M \times N$, its MSCN coefficients are calculated by:

$$\widehat{I}(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + 1}, \quad i \in \{1,2,\cdots,M\}, \quad j \in \{1,2,\cdots,N\}, \tag{1}$$

with

$$\mu(i,j) = \sum_{h=-H}^{h=H} \sum_{r=-R}^{r=R} \omega_{h,r} \cdot I_{h,r}(i,j) \tag{2}$$

$$\sigma(i,j) = \sqrt{\sum_{h=-H}^{h=H} \sum_{r=-R}^{r=R} \omega_{h,r} \cdot (I_{h,r}(i,j) - \mu(i,j))^2}, \tag{3}$$

where $I(i,j)$ denotes the pixel value in the (i,j)th spatial location; $\omega = \{\omega_{h,r} | h = -H, \cdots, H; r = -R, \cdots, R\}$ defines a 2D weighting filter. H and R determine the size of local patch considered in the calculation of the mean $\mu(i,j)$ and standard deviation $\sigma(i,j)$. Actually, the distribution of MSCN coefficients $\widehat{I}$ is not sensitive to filter size or type [23]. Hence, this study adopts the same 2D Gaussian kernel with the size of 7×7 (i.e., $H = R = 3$) as in [23].

As reported by Ruderman [42], the MSCN coefficients of a certain natural image interestingly exhibit Gaussian-like distribution, while distortion modifies the distribution towards Laplacian or Weibull one. In this sense, the distortion level of a natural image is closely related to its corresponding MSCN coefficients' distribution statistics which can be well estimated by GGD. As MSCN coefficients distribute symmetrically, a zero mean GGD defined by Eq. (4) is utilized:

$$f(x,\alpha,\nu^2) = \frac{\alpha}{2\beta\Gamma(1/\alpha)} exp\big(-(|x|/\beta)^2\big), \tag{4}$$

where β is an intermediate variable,

$$\beta = \nu\sqrt{\Gamma(1/\alpha)/\Gamma(3/\alpha)}, \tag{5}$$

with gamma function $\Gamma(\cdot)$ defined as:

$$\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt, \quad a > 0. \tag{6}$$

In Eq. (4), α and ν^2, reflecting the image naturalness, control the shape and variance of distribution, respectively. x is the MSCN coefficient. In this study, the GGD model is deployed to fit MSCN distributions of left and right views, and the parameters α and ν^2 are regarded as the quality-sensitive features $\boldsymbol{f}_1$.

As human visual system (HVS) adapts to the mean background value, logarithmic intensity functions are used to remove background [42]. Formally, the 'log-contrast' function $\psi(i,j)$ is defined as:

$$\psi(i,j) = ln(I(i,j) + \epsilon) - ln(I_0), \tag{7}$$

where, $\epsilon = 0.01$ is used to avoid infinite value; I_0 is defined satisfying $\Sigma\psi(i,j) = 0$. Then, the log-contrast value $\psi(i,j)$ is normalized with respect to its local standard deviations.

$$\varphi(i,j) = \frac{\psi(i,j) - \overline{\psi}(i,j)}{\rho(i,j) + C_1}, \tag{8}$$

where $\overline{\psi}(i,j)$ and $\rho(i,j)$ calculated with Eqs. (2) and (3) are the mean and standard deviation of $\psi(i,j)$; $C_1 = 0.01$ is a constant that prevents instabilities when the denominator tends to zero. Operated with Eq. (8), log-variance normalized image $\varphi(i,j)$ tends to be more uniform than the original image and seems to show a Gaussian-like distribution, as depicted in Fig. 2. Besides, the standard deviation image $\rho(i,j)$ looks like the original image to some extent. Therefore, taking $\rho(i,j)$ as an original image, its log-contrast is further expressed as:

$$\xi(i,j) = ln(\rho^2(i,j) + \epsilon) - ln(\rho_0), \tag{9}$$

where ρ_0 is analogous to I_0. Moreover, the statistics of ξ are found to be similar to those of φ obtained by Eq. (8). The above operations can be further re-applied in multiple scales by taking ξ as the input. To be more specific, ξ is first operated by Eq. (9), and the output is fed into Eq. (8) to obtain the variance normalized image and standard deviation image, which is taken as the input of next iteration. Fig. 3 depicts the outline of MSCN coefficients' generation in a multi-scale manner. It is worth emphasizing that the MSCN coefficients of the log-variance normalized images show Gaussian-like distribution. Therefore, the GGD parameters fitted on the MSCN coefficients of such multi-scale log-variance are extracted as the quality-sensitive features $\boldsymbol{f}_2$.

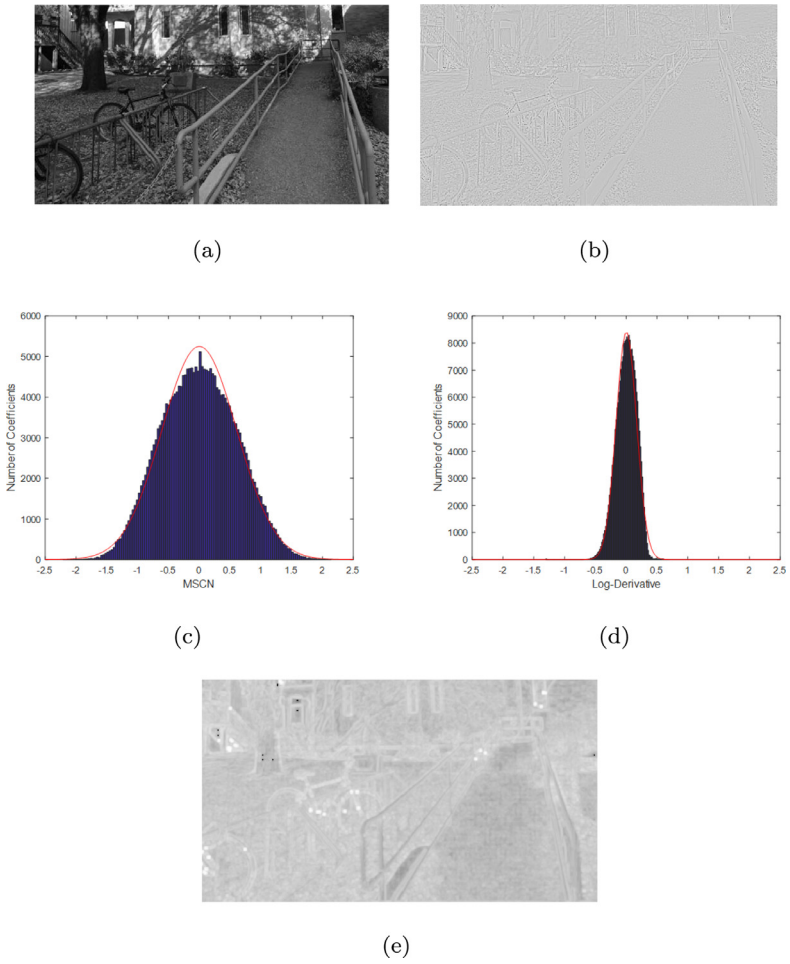

Fig. 2. Image visualization, (a) and (b) are original image and its log-variance normalized image, while (c) and (d) are their associated distributions of MSCN coefficients. (e) is the standard deviation map of (b).

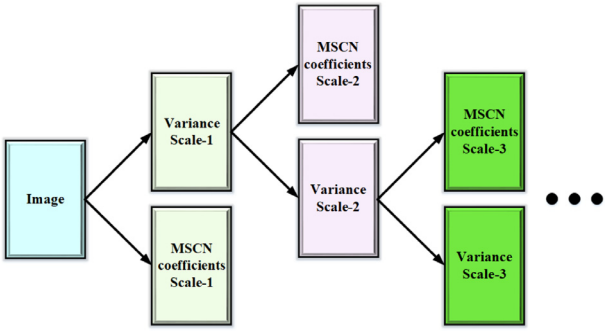

Fig. 3. Multiscale generation of the MSCN coefficients by iterating over the variance images.

2.2. Cyclopean map and its associated features

One of the key aspects of the visual cortex is binocularity. The monocular stimuli to the left and right eyes first pass through an interocular gain-control pathway, and then combine to form a single cyclopean perception. In the literature, cyclopean image, providing an approximation of the perceived image in the brain from the input left and right views, is simply synthesized by a linear model [14]:

$$Cyc(x,y) = W_L(x,y)\cdot I_L(x,y) + W_R(x,y)\cdot \tilde{I}_R(x,y), \tag{10}$$

where I_L and I_R are the left and right views, respectively; $\tilde{I}_R(x,y) = I_R(x+d(x,y),y)$ is the disparity-compensated view; $d(x,y)$ is the horizontal disparity value in the (x,y)th location. In this paper, the disparity value is calculated by SSIM-based algorithm utilized in [14]. $W_L(x,y)$ and $W_R(x,y)$ are corresponding weights obtained by a gain-control model [43]:

$$W_L(x,y) = \frac{E_L(x,y)}{E_L(x,y)+\tilde{E}_R(x,y)} \tag{11}$$

$$W_R(x,y) = \frac{\tilde{E}_R(x,y)}{E_L(x,y)+\tilde{E}_R(x,y)}, \tag{12}$$

where E_L and $\tilde{E}_R$ are 2D Gabor filter magnitude responses of the left and disparity-compensated views, respectively. Actually, various parameters are involved in Gabor function. It is out of scope to study the impacts of these parameters on the Gabor function in this work. This paper sets these parameters by strictly referring to reference [44]. Fig. 4(a)–(c) give the left and right views and the corresponding cyclopean view. As seen, the cyclopean view looks like the left or right view and also exhibits 'naturalness' to some extent. Accordingly, its MSCN coefficients also show a Gaussian-like distribution, as shown in Fig. 4(f). Therefore, the GGD parameters of its fitted distribution are also extracted as the quality-sensitive features $\boldsymbol{f}_3$.

Because retinal and cortical neurons in the visual cortex respond selectively to stimulus orientation and frequency, the statistics of multi-scale and multi-orientation filter responses resemble

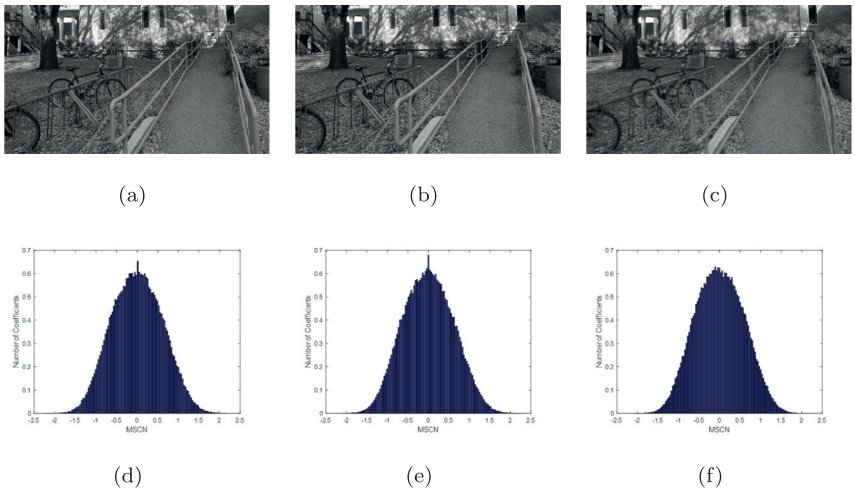

Fig. 4. Examples of left, right views and generated cyclopean map, (a) left view, (b) right view and (c) cyclopean map. (d)–(f) are corresponding distributions of MSCN coefficients.

the diverse scale and orientation sensitivities of the receptive fields of retinal and cortical neurons [45]. In this paper, the 2D log-Gabor filters are further deployed to extract perceptually relevant features on cyclopean view. Formally, it is expressed as:

$$G_{s,o}(\varpi,\theta) = exp\left[-\left(\frac{log\left(\frac{\varpi}{\varpi_s}\right)}{2\delta_s^2}\right)\right]\cdot exp\left[-\left(\frac{(\theta-\theta_o)^2}{2\delta_o^2}\right)\right], \quad (13)$$

where ϖ and θ are the normalized radial frequency and the orientation angle of the filter; ϖ_s and θ_o are the corresponding center frequency and orientation of the filter, respectively; δ_s and δ_o are the constants to determine the strength of the filter. All these parameters are strictly arranged following reference [36].

Given that phase information has stronger ability of reflecting image distortion than the amplitude information [46], features are also extracted on the cyclopean phase map, which is defined as:

$$\kappa(x,y) = arctan\left(\frac{\zeta(x,y)}{\eta(x,y)}\right), \quad (14)$$

where ζ denotes the mean value of imaginary part of log-Gabor responses along 4 scales and 6 orientations, and η corresponds to that of real part. According to Marr's theory [47], the structural information, which forms primitives in V1 area (the first visual perception cortical area), is closely related to the perceptual quality. An effective IQA metric may benefit from estimating the structural information modification. Gray level co-occurrence matrix (GLCM), calculating the statistical characteristics between image pixels, has been proved that it can effectively reflect pixel value variation [48]. Therefore, GLCM is employed to extract structural information. The GLCM is made up of the joint probability densities between the gray levels of the image. Given an image, its GLCM with size $L\times L$ can be obtained by analyzing occurrences of the gray pair [48], where L is the arranged number of gray level (for more details about GLCM, please refer to [48]). To date, numerous features have been extracted to analyze image properties from GLCM map [48]. In this work, two features, namely energy and homogeneity, are selected for IQA. Specifically, energy describes the distribution and roughness of overall GLCM. A small energy value indicates that the image possesses uniform distribution. Formally, it can be defined as:

$$E = \sum_{i=1}^{L}\sum_{j=1}^{L} P_{D,\Theta}(i,j)^2, \quad (15)$$

where $P_{D,\Theta}(i,j)$ stands for the probability of an entry $m_{D,\Theta}(i,j)$ in the co-occurrence matrix $M_{D,\Theta}$. D and Θ are the distance and orientation considered in joint probability calculation. $P_{D,\Theta}(i,j)$ can be expressed as:

$$P_{D,\Theta}(i,j) = \frac{m_{D,\Theta}(i,j)}{\sum_{i=1}^{L}\sum_{j=1}^{L} m_{D,\Theta}(i,j)}. \quad (16)$$

Homogeneity measures local change of GLCM by capturing local gray correlation. More concretely, a large value of homogeneity corresponds to small local change of image. By definition, homogeneity can be computed as:

$$H = \frac{P_{D,\Theta}(i,j)}{|i-j-2|+1}. \quad (17)$$

In this paper, D is set as 1 and features are extracted in four directions (i.e., $\Theta = \{0°, 45°, 90°, 135°\}$). All these GLCM-based features are adopted for IQA and denoted as $\boldsymbol{f}_4$.

2.3. Distortion asymmetry

S3D image contains two views and is easily subject to asymmetric distortions, i.e., distortion levels of the left and right views are different. Compared to symmetric distortion, evaluating the asymmetric distortion is more challenging. Here, a simple method is proposed by directly calculating the difference map between left and right views to quantify the asymmetry. Fig. 5 shows the difference map between Fig. 4(a) and (b) as well as the distribution of MSCN coefficients. It is clear that the histogram of difference map exhibits Gaussian-like distribution, as shown in Fig. 5(b). When suffered from asymmetric distortion, the difference map also exhibits Gaussian-like distribution but with shape change. To verify this phenomenon, an experiment is conducted by keeping the left view distortion-free and blurring the right view with different levels. Fig. 5(c) depicts the distribution of MSCN coefficients under three situations. It is intuitive that the distribution shape (especially the peak value) changes when increasing the asymmetric blurriness, which indicates that the GGD parameters of difference map can be used to reflect the asymmetric degree. Moreover, considering that HVS is more sensitive to structural information, therefore, the gradient similarity between left and right view is also computed as a feature to reflect the distortion asymmetry. The parameters of GGD fit and gradient similarity are combined as quality-sensitive features $\boldsymbol{f}_5$.

So far, multiple features ($\mathbf{F} = [\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3, \boldsymbol{f}_4, \boldsymbol{f}_5]$) have been obtained by comprehensively considering different S3D quality-

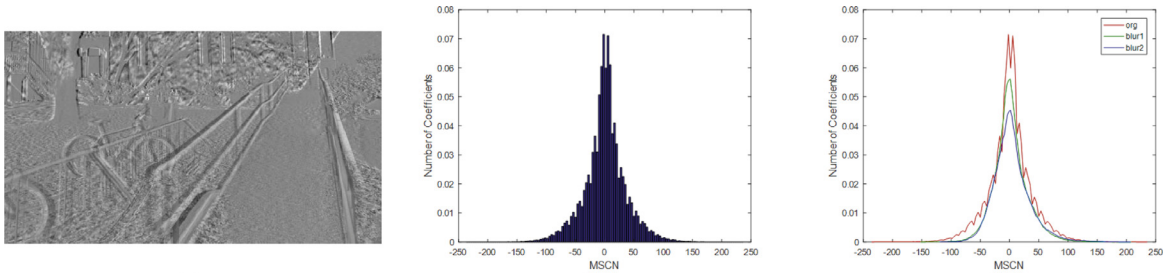

Fig. 5. (a) is the difference map between Fig. 4(a) and (b); (b) is the associated distribution of MSCN coefficients; (c) is the distribution map under different asymmetric distortions. 'org' represents the original image, while 'blur1' and 'blur2' denote two blurred versions (blurring degree: 'blur1' < 'blur2').

relevant factors, i.e., naturalness, structure and distortion asymmetry. Furthermore, the image is down-sampled into multi-scale and these features are extracted on each scale in view of perceivability of image details as [49]. To help readers better understand the whole process, the pseudo-code of the feature extraction procedure is given in Algorithm 1.

Algorithm 1 Procedure of feature extraction.

Input: Stereopair, i.e., left view I_L and right view I_R;
Output: Feature vector, **F**;
1: Extracting the GGD parameters ($\boldsymbol{f}_1$) of MSCN coefficients of I_L and I_R;
2: Extracting the GGD parameters ($\boldsymbol{f}_2$) of MSCN coefficients of log-variance normalized images of I_L and I_R;
3: Calculating the disparity map d using stereo match algorithm and generating the cyclopean map Cyc using I_L, I_R and d;
4: Extracting the GGD parameters ($\boldsymbol{f}_3$) of MSCN coefficients of Cyc;
5: Extracting features ($\boldsymbol{f}_4$) from the GLCM matrix of the phase map of Cyc;
6: Extracting the GGD parameters of MSCN coefficients of the difference map between I_L and I_R; Computing the similarity between I_L and I_R; Combining these two parts' features as ($\boldsymbol{f}_5$);
7: Integrating feature vector $\mathbf{F} = [\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3, \boldsymbol{f}_4, \boldsymbol{f}_5]$;
8: **return F**;

2.4. Image quality estimation

After obtaining quality-sensitive features, how to utilize these features for quality estimation is the next problem. Machine learning, transforming high-dimensional features into a lower-dimensional representation, has been proved as an effective tool for solving prediction or nonlinear fitting problem and it is therefore utilized for image quality estimation in this study. This process contains two stages, training stage and testing stage, as shown in Fig. 1. During training stage, quality-sensitive features of training database are first extracted and integrated into basic features as training feature vector F_{tr}. Then, a regression algorithm is applied to learn a prediction function f_p from F_{tr} to subjective ratings (e.g., MOS or DMOS). During testing stage, the feature vector F_{te} is first extracted, image quality scores of testing images are subsequently predicted by feeding F_{te} into the well trained f_p. Recently, SVR has been successfully used in IQA. In this paper, the Library for Support Vector Machines (LIBSVM) tool is chosen to implement ϵ-SVR [50]. As radial basis function (RBF) has fast convergence characteristic and approximates to various nonlinear functions, it is chosen as the kernel function of ϵ-SVR to measure the similarity between two samples in high dimension.

3. Experimental results and analysis

3.1. Experiment setup

3.1.1. S3D IQA database

In order to verify the effectiveness and make comparisons with mainstream IQA metrics, four publicly available S3D IQA databases are chosen as the test platforms.

(1) **LIVE 3D IQA Phase I database** [51]: This database totally consists of 365 symmetrically distorted images corrupted from 20 reference images. It includes five distortion types, i.e., JPEG, JPEG2000 (JP2K), Gaussian blur (Gblur), White noise (WN) as well as the simulated fast-fading Rayleigh channel (FF), with various quantities. The subjective ratings are reported in forms of DMOS.
(2) **LIVE 3D IQA Phase II database** [14]: Compared with LIVE 3D IQA Phase I database, LIVE 3D IQA Phase II database is more complex with both symmetric and asymmetric distortions. To be specific, this database totally contains 360 distorted images (with 5 distortion types as same as in the LIVE 3D IQA Phase I database) corrupted from 8 reference images. For each distortion type, every reference image is processed to create 9 distorted images with 3 symmetrically distorted versions and 6 asymmetrically distorted versions. Therefore, 72 distorted images make up of each distortion type. Each image is assigned with subjective ratings in forms of DMOS.
(3) **Waterloo-IVC 3D IQA Phase I database** [28]: This database contains both 2D and 3D images. Only 3D images are selected for the sake of the purpose of this work. Totally, it consists of 330 symmetrically distorted images corrupted from 6 reference images. Each reference image is processed by three distortion types, i.e., WN, JPEG and Gblur.
(4) **Waterloo-IVC 3D IQA Phase II database** [28]: This database contains both symmetrically and asymmetrically distorted 3D images corrupted from 10 reference images. It is the first database that contains mixed distortion types in asymmetrically distorted images. For each distorted image, its quality is given in form of MOS.

3.1.2. Performance indicators

In this study, three commonly used criteria are utilized to quantify and verify the performance of the proposed method as well as all competing metrics (as described in next subsection). They are Spearman rank-order correlation coefficient (SRCC), Pearson linear correlation coefficient (PLCC), and Root mean square error (RMSE). Among them, SRCC and PLCC evaluate prediction monotonicity and consistency of prediction performance, respectively. Whereas, RMSE reflects the prediction error. For a perfect metric, the match value between the predicted scores and associated subjective ratings is close to 1 for PLCC and SRCC but close to 0 for RMSE. In

Table 1
Results of the proposed method under different settings.

Database	Criteria	Log-derivative Scale					Down-sample Scale		
		1	2	3	4	5	1	2	3
LIVE I	SRCC	0.884	0.898	0.908	0.914	0.910	0.903	0.914	0.904
LIVE II	SRCC	0.882	0.894	0.903	0.906	0.905	0.885	0.906	0.904
LIVE I	PLCC	0.900	0.907	0.915	0.937	0.924	0.910	0.937	0.929
LIVE II	PLCC	0.890	0.897	0.907	0.914	0.912	0.899	0.914	0.906
LIVE I	RMSE	6.187	6.002	5.796	5.652	5.723	5.961	5.652	5.982
LIVE II	RMSE	4.712	4.603	4.562	4.445	4.524	4.753	4.445	4.781

Table 2
Performance comparisons on LIVE databases among the proposed method and competing FR methods.[a]

Database	Criteria	2D IQA Metrics		3D IQA Metrics				
		VSI	SSIM	M2	FI-PSNR	M1	M3	Pro. Method
LIVE I	PLCC	0.863	0.877	0.866	0.902	0.924	0.931	**0.937**
	SRCC	0.865	0.877	0.856	0.899	0.916	**0.933**	0.914
	RMSE	8.288	7.879	8.201	7.061	6.268	6.057	**5.652**
LIVE II	PLCC	0.766	0.803	0.662	0.748	0.907	0.822	**0.914**
	SRCC	0.748	0.792	0.638	0.728	0.901	0.804	**0.906**
	RMSE	7.262	6.727	8.465	7.490	4.746	6.574	**4.449**

[a] Due to the space limitation, 'M1', 'M2', and 'M3' are used to represent Chen's method [14], Benoit's method [52], and Shao's method [29], respectively.

order to reduce the nonlinearity, a five-parameter logistic regression function is applied before the computation of PLCC and RMSE. Formally, it can be express as:

$$g(q) = \tau_1 \cdot \left(\frac{1}{2} - \frac{1}{exp(\tau_2 \cdot (q - \tau_3)) + 1}\right) + \tau_4 \cdot q + \tau_5 \tag{18}$$

where $g(q)$ is the fitted quality score, q denotes the objectively predicted score. τ_ϑ, $\vartheta = \{1, 2, \cdots, 5\}$ is the parameter vector to be fitted.

3.1.3. Parameter settings

Before conducting comparisons, two parameters, i.e., log-derivative scale and down-sample scale, should be determined in advance. In this study, they are determined based on the experiment on LIVE databases. Taking down-sample scale as the breakthrough point (i.e., fix it as 2), Table 1 illustrates experimental results when the log-derivative scale is respectively set as 1, 2, 3, 4 and 5 on the left part. From the table, it is clear that: first, only moderate performance is obtained when log-derivative scale is small (e.g., 1, 2); second, the performance increases with the growth of log-derivative scale and reaches the peak (when log-derivative scale is 4). Hence, the log-derivative scale is set as 4 in this study. Likewise, by fixing the log-derivative scale as 4, the influence of down-sample scale is given on the right side of Table 1 when varying down-sample scale from 1 to 3. Obviously, similar results appear as compared to those obtained in log-derivative scale determination experiment. With these observations, the log-derivative scale and down-sample scale are set as 4 and 2, respectively.

3.2. Overall performance comparison

The proposed method is compared against existing mainstream IQA methods, including six FR metrics (SSIM [2], VSI [18], Chen's method [14], FI-PSNR [15], Benoit's method [52], and Shao's method [29]), and seven NR metrics (NIQE [53], BLIINDS-II [22], StereoQUE [54], Shao's method [55], Zhou's method [35], Zhou's method [36], and Chen's method [56]). Note that for 2D IQA metrics (i.e., SSIM, VSI, NIQE, and BLIINDS-II), the scores of left and right views are calculated separately and the overall score of S3D image is obtained as their mean value. In the implementation of learning-based metrics, the dataset requires to be randomly split into two non-overlapping subsets: training subset and testing subset. This study strictly follows previous works and takes 80% dataset as training subset, while the resting is regarded as testing subset. After training the model on the training subset, the prediction performance is measured on the test subset. To avoid the performance bias, the random training-testing split is repeated 1000 times and the performance is reported in form of median value.

Tables 2–4 tabulate the experimental results obtained through the implementation of the released demos on four databases. As some metrics' source codes (e.g., Benoit's method [52], StereoQUE [54], Shao's method [29], Shao's method [55], Zhou's method [35], and Zhou's method [36]) are not available, those associated results are directly excerpted from original papers [29,35,36,52,54,55]. For results that are not given, the associated comparisons are discarded and the tables are filled with symbol '-'. (Since some works did not report results on Waterloo-IVC databases, the comparison results are specially listed without considering such works on Waterloo-IVC databases.) The best performance is highlighted in boldface for readers' convenience. From the experimental results, it is evident that the proposed method is superior to all competing methods. Such conclusion can be drawn from the analyses of the following aspects: (1) Several competing 2D FR and NR IQA metrics, i.e., SSIM, VSI, NIQE, and BLIINDS-II, are not competent to solve S3D IQA problem and are inferior to the proposed method. This possibly attributes to that 2D metrics ignore the depth information. (2) The proposed method, together with most S3D FR IQA metrics, provides objectively and considerably accurate quality predictions. Moreover, it outperforms all competing 3D metrics on LIVE II, Waterloo-IVC I-II databases and is comparative to 'M3' (Shao' metric [29]) on LIVE I database. These encouraging performance benefits from sufficiently considering binocular visual perception, where depth information occupies an important role. (3) Besides, the proposed method outperforms all competing metrics in terms of stability. Specifically, the vast majority of competing metrics provide distinct performance on symmetric and asymmetric distortions. To specify, they obtain general performance on sym-

Table 3
Performance comparisons on LIVE databases among the proposed method and competing NR methods.[a]

Database	Criteria	2D IQA Metrics		3D IQA Metrics					
		NIQE	BLIINDS-II	M7	M6	StereoQUE	M4	M5	Pro. Method
LIVE I	PLCC	0.812	0.823	0.895	0.907	0.917	0.928	0.929	**0.937**
	SRCC	0.801	0.821	0.891	0.896	0.911	0.887	0.901	**0.914**
	RMSE	8.854	9.366	7.247	–	6.598	6.025	6.010	**5.652**
LIVE II	PLCC	0.740	0.907	0.895	0.848	0.845	0.861	0.856	**0.914**
	SRCC	0.708	0.901	0.880	0.824	0.888	0.823	0.819	**0.906**
	RMSE	8.855	4.746	5.102	–	7.297	5.779	6.041	**4.449**

[a] Due to the space limitation, 'M4', 'M5', 'M6', and 'M7' are used to represent Zhou's method [35], Zhou's method [36], Shao's method [55], and Chen's method [56], respectively.

Table 4
Performance comparisons on Waterloo-IVC databases among the proposed method and competing methods.

Database	Criteria	FR IQA Metrics				NR IQA Metrics		
		SSIM	VSI	M1	FI-PSNR	NIQE	BLIINDS-II	Pro. Method
Waterloo-IVC I	PLCC	0.634	0.789	0.639	0.665	0.510	0.770	**0.926**
	SRCC	0.490	0.737	0.489	0.499	0.229	0.710	**0.919**
	RMSE	12.164	9.664	12.111	11.752	13.535	10.044	**4.610**
Waterloo-IVC II	PLCC	0.484	0.644	0.543	0.202	0.124	0.729	**0.911**
	SRCC	0.389	0.601	0.360	0.369	0.157	0.681	**0.895**
	RMSE	16.749	14.556	16.075	18.745	18.993	13.105	**7.737**

Table 5
Influence of asymmetry-sensitive features.

Database	With			Without		
	PLCC	SRCC	RMSE	PLCC	SRCC	RMSE
LIVE I	0.937	0.914	5.652	0.929	0.910	5.976
LIVE II	0.914	0.906	4.449	0.908	0.893	4.623
Waterloo-IVC I	0.926	0.919	4.610	0.921	0.916	4.752
Waterloo-IVC II	0.911	0.895	7.737	0.900	0.889	7.949

metric distortion but fail in asymmetric distortion. While the proposed method achieves a reasonable performance balance across all databases with PLCC $>$ 0.9. A reasonable explanation is that the proposed method captures the asymmetric features by excavating the difference between left and right views. From the above analysis, the proposed metric objectively exhibits a powerful potential and ability on perceptual quality assessment of S3D images.

As stated above, the proposed method can effectively handle asymmetric distortion quality assessment problem. One reason is that the proposed method utilizes asymmetry-sensitive features, i.e., the GGD fit parameters of difference map between left and right views and the similarity between their gradient maps. Here, an extra experiment is further conducted to explore the influence of these asymmetry-sensitive features. Table 5 tabulates the results of the proposed method with/without considering these features. Obviously, no significant bias happens on symmetric distorted databases, such as LIVE I and Waterloo-IVC I. However, it is worth emphasizing that experimental results obviously change on asymmetrically distorted databases. Specifically, without the help of these features, the PLCC and SRCC performances have dropped about 1%. Based on the above analysis, it can be concluded that the proposed method possesses advantages on quality assessment of asymmetric distortions.

For learning-based metrics, the ratio between training subset and testing subset sensitively affects the experimental performance. Ideally, an excellent metric's performance is stable with a little perturbation, ignoring the ratio change. For this purpose, multiple ratios are arranged by varying the training subset from 30% to 90% in steps of 10% and taking the rest as testing subset. All results are reported in form of median from 1000 randomly train-test splits. Fig. 6 depicts the relationship between SRCC and

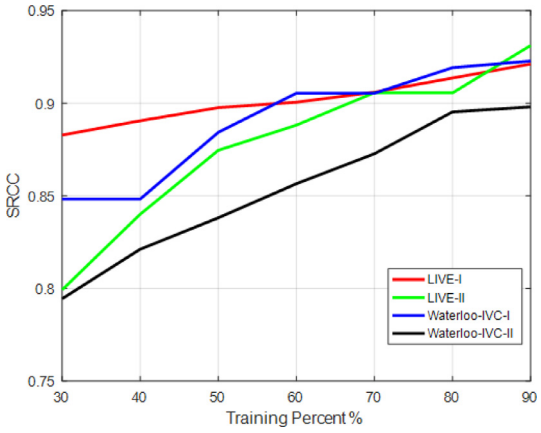

Fig. 6. SRCC values under different training subsets.

training ratio. Obviously, the performance increases with the increasing of training percent and is still acceptable even though the training set is less than 50%.

3.3. Performance on individual distortion types

An excellent IQA metric should not only show its powerful ability on entire database, but also effectively cope with quality assessment task on each individual distortion type. In this section, an experiment is further conducted to investigate and verify the superiority of the proposed method on estimating the perceived quality of individual distortion type. Given that Waterloo-IVC databases are more complex and contain mixed distortion types, therefore, they are not suitable for this task. Tables 6 and 7 tabulate the experimental results. Similar to Table 3, the best result for NR IQA comparisons is also highlighted in bold. Besides, the average performance ranking of all comparisons on each distortion type is given in the last column. As seen, almost all metrics do not work well on all individual distortions but only show effectiveness on some distortion types. Compared with FR metrics, NR metrics obtain quite different rank results across different databases. Generally, it is unfair for NR metric to compare with FR metric, as NR metric lacks the reference information. In spite of this, the pro-

Table 6
Performance comparisons among different IQA metrics on individual distortion types (SRCC).

Database	Metrics	Type	JPEG	JP2K	Gblur	WN	FF	Rank
LIVE I	SSIM	FR	0.435	0.858	0.879	0.939	0.587	5
	VSI	FR	0.428	0.864	0.895	0.938	0.618	4
	FI-PSNR	FR	0.207	0.839	0.935	0.928	0.658	6
	M1	FR	0.558	0.896	0.926	**0.948**	0.688	3
	M2	FR	**0.603**	**0.910**	0.931	0.930	0.699	2
	M3	FR	0.495	0.895	**0.941**	0.940	**0.796**	1
	NIQE	NR	0.599	0.632	0.861	0.907	0.519	8
	BLIINDS-II	NR	**0.801**	0.905	0.821	0.867	**0.896**	3
	StereoQUE	NR	0.782	**0.917**	0.865	0.910	0.666	2
	M4	NR	0.614	0.824	0.916	0.915	0.867	1
	M5	NR	0.562	0.856	0.897	0.921	0.771	4
	M6	NR	0.427	0.871	**0.932**	0.914	–	5
	M7	NR	0.617	0.863	0.878	0.919	0.652	7
	Pro. Method	NR	0.595	0.832	0.857	**0.932**	0.779	6
LIVE II	SSIM	FR	0.678	0.704	0.838	0.922	0.834	5
	VSI	FR	0.716	0.655	0.727	**0.957**	0.790	4
	FI-PSNR	FR	0.613	0.719	0.711	0.907	0.701	6
	M1	FR	0.840	**0.833**	0.910	0.955	0.889	1
	M2	FR	**0.867**	0.751	0.455	0.923	0.773	3
	M3	FR	0.733	0.785	**0.965**	0.920	**0.890**	2
	NIQE	NR	0.637	0.599	0.851	0.775	0.600	7
	BLIINDS-II	NR	0.718	0.579	0.815	0.785	0.840	5
	StereoQUE	NR	0.839	0.864	0.846	0.932	0.888	4
	M4	NR	0.593	0.717	0.903	0.891	0.891	3
	M5	NR	0.737	0.647	**0.911**	0.936	0.798	6
	M6	NR	–	–	–	–	–	–
	M7	NR	**0.867**	0.867	0.900	0.950	**0.933**	1
	Pro. Method	NR	0.769	**0.959**	0.868	**0.959**	0.913	2

Table 7
Performance comparisons among different IQA metrics on individual distortion types (PLCC).

Database	Metrics	Type	JPEG	JP2K	Gblur	WN	FF	Rank
LIVE I	SSIM	FR	0.488	0.875	0.919	0.944	0.724	5
	VSI	FR	0.498	0.876	0.819	0.945	0.739	4
	FI-PSNR	FR	0.265	0.839	0.811	0.927	0.710	6
	M1	FR	0.634	0.929	0.811	**0.957**	0.761	3
	M2	FR	**0.640**	**0.939**	0.948	0.925	0.747	2
	M3	FR	0.498	0.921	**0.959**	0.944	**0.859**	1
	NIQE	NR	0.589	0.748	0.919	0.908	0.579	5
	BLIINDS-II	NR	0.730	0.757	0.864	0.905	0.714	6
	StereoQUE	NR	**0.806**	**0.938**	0.881	0.919	0.758	2
	M4	NR	–	–	–	–	–	–
	M5	NR	–	–	–	–	–	–
	M6	NR	0.458	0.901	0.916	0.952	–	4
	M7	NR	0.695	0.907	0.917	0.917	0.735	3
	Pro. Method	NR	0.744	0.934	**0.971**	**0.962**	**0.854**	1
LIVE II	SSIM	FR	0.666	0.726	0.849	0.931	0.861	3
	VSI	FR	0.699	0.684	0.777	0.958	0.841	6
	FI-PSNR	FR	0.628	0.755	0.739	0.915	0.705	4
	M1	FR	**0.843**	**0.844**	**0.966**	**0.962**	**0.916**	1
	M2	FR	0.533	0.647	0.887	0.861	0.847	5
	M3	FR	0.747	0.782	0.958	0.946	0.905	2
	NIQE	NR	0.712	0.654	0.969	0.604	0.760	5
	BLIINDS-II	NR	0.767	0.579	0.920	0.785	0.879	6
	StereoQUE	NR	0.829	0.867	0.878	0.920	0.836	4
	M4	NR	–	–	–	–	–	–
	M5	NR	–	–	–	–	–	–
	M6	NR	0.827	0.826	0.928	0.984	–	3
	M7	NR	**0.867**	0.867	0.900	0.950	**0.933**	2
	Pro. Method	NR	0.843	**0.986**	**0.973**	**0.986**	0.923	1

posed method is still comparable with FR metrics on some distortion types, e.g., WN and FF. Besides, the proposed method shows powerful competitiveness on LIVE II database. Specifically, it totally obtains the best result 2 times and occupies the top three places 4 times on total 5 SRCC comparisons. Meanwhile, the proposed method always occupies the top place on the average PLCC rank, exhibiting its superiority incisively and vividly. To this end, these observations verify that the proposed method can be fully competent to tackle quality assessment problem of S3D images.

4. Conclusion and future work

In this paper, an effective blind quality assessment algorithm, which can manage both symmetric and asymmetric distortions, has been introduced for S3D images. First, a number of quality-sensitive features are captured to represent image distortions. More specifically, NSS features are extracted on both left and right views to reflect the image naturalness. Whereas, statistical features are captured on cyclopean map, which is created by consider-

ing binocular properties. Besides, to quantify the asymmetric distortion, the distribution of difference map between left and right views is fitted and the similarity between them is computed. Finally, a regression model is learnt based on SVR to estimate the quality of a test image. Experimental results prove its superiority on four publicly available databases.

In the future, some works can be conducted on the following aspects to further promote the development of this research topic: (1) building a sufficiently large-scale S3D image database by considering both visual discomfort and multiple distortions; (2) integrating visual discomfort factor with image quality to evaluate the complete quality of experience; (3) extending this work to solve quality assessment of S3D videos.

Acknowledgments

Support for this work is provided by the National Natural Science Foundation of China under Grants 61520106002, 61731003 and 61471262. This work is also supported by China Scholarship Council (Grant no. 201706250189).

References

[1] W. Lin, C.-C.J. Kuo, Perceptual visual quality metrics: a survey, J. Vis. Commun. Image Represent. 22 (4) (2011) 297–312.
[2] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[3] Y. Yang, J. Ming, Image quality assessment based on the space similarity decomposition model, Signal Process. 120 (2016) 797–805.
[4] X. Min, K. Gu, G. Zhai, M. Hu, X. Yang, Saliency-induced reduced-reference quality index for natural scene and screen content images, Signal Process. 145 (2018) 127–136.
[5] K. Gu, J. Qiao, X. Min, G. Yue, W. Lin, D. Thalmann, Evaluating quality of screen content images via structural variation analysis, IEEE Trans. Vis. Comput. Graph. (2017), doi:10.1109/TVCG.2017.2771284.
[6] G. Yue, C. Hou, K. Gu, N. Ling, B. Li, Analysis of structural characteristics for quality assessment of multiply distorted images, IEEE Trans. Multim. (2018), doi:10.1109/TMM.2018.2807589.
[7] A. Saha, Q.J. Wu, Full-reference image quality assessment by combining global and local distortion measures, Signal Process. 128 (2016) 186–197.
[8] Z. Shi, J. Zhang, Q. Cao, K. Pang, T. Luo, Full-reference image quality assessment based on image segmentation with edge feature, Signal Process. 145 (2018) 99–105.
[9] J. You, L. Xing, A. Perkis, X. Wang, Perceptual quality assessment for stereoscopic images based on 2D image quality metrics and disparity analysis, in: Proc. of International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, AZ, USA, 2010.
[10] F. Shao, W. Lin, Z. Li, G. Jiang, Q. Dai, Toward simultaneous visual comfort and depth sensation optimization for stereoscopic 3-D experience, IEEE Trans. Cybern. (2016), doi:10.1109/TCYB.2016.2615856.
[11] G. Yue, C. Hou, K. Lu, D. Feng, Y. Li, Subjective visual comfort assessment based on fusion time for depth information, in: Computer Science & Education (ICCSE), 2016 11th International Conference on, IEEE, 2016, pp. 733–737.
[12] G. Yue, C. Hou, J. Lei, Y. Fang, W. Lin, Optimal region selection for stereoscopic video subtitle insertion, IEEE Trans. Circuits Syst. Video Technol. (2017,), doi:10.1109/TCSVT.2017.2739756.
[13] F. Shao, W. Tian, W. Lin, G. Jiang, Q. Dai, Learning sparse representation for no-reference quality assessment of multiply-distorted stereoscopic images, IEEE Trans. Multim. 18 (10) (2017) 2104–2114.
[14] M.-J. Chen, C.-C. Su, D.-K. Kwon, L.K. Cormack, A.C. Bovik, Full-reference quality assessment of stereopairs accounting for rivalry, Signal Process. Image Commun. 28 (9) (2013) 1143–1155.
[15] Y.-H. Lin, J.-L. Wu, Quality assessment of stereoscopic 3D image compression by binocular integration behaviors, IEEE Trans. Image Process. 23 (4) (2014) 1527–1542.
[16] R. Bensalma, M.-C. Larabi, A perceptual metric for stereoscopic image quality assessment based on the binocular energy, Multidimens. Syst. Signal Process. 24 (2) (2013) 281–316.
[17] G. Yue, C. Hou, K. Gu, Subjective quality assessment of animation images, in: Visual Communications and Image Processing (VCIP), 2017 IEEE, IEEE, 2017, pp. 1–4.
[18] L. Zhang, Y. Shen, H. Li, VSI: A visual saliency-induced index for perceptual image quality assessment, IEEE Trans. Image Process. 23 (10) (2014) 4270–4281.
[19] D.M. Chandler, S.S. Hemami, VSNR: A wavelet-based visual signal-to-noise ratio for natural images, IEEE Trans. Image Process. 16 (9) (2007) 2284–2298.
[20] K. Gu, S. Wang, G. Zhai, W. Lin, X. Yang, W. Zhang, Analysis of distortion distribution for pooling in image quality prediction, IEEE Trans. Broadcast. 62 (2) (2016) 446–456.
[21] S. Wang, K. Gu, X. Zhang, W. Lin, S. Ma, W. Gao, Reduced-reference quality assessment of screen content images, IEEE Trans. Circuits Syst. Video Technol. 28 (1) (2018) 1–14.
[22] M.A. Saad, A.C. Bovik, C. Charrier, Blind image quality assessment: a natural scene statistics approach in the DCT domain, IEEE Trans. Image Process. 21 (8) (2012) 3339–3352.
[23] A. Mittal, A.K. Moorthy, A.C. Bovik, No-reference image quality assessment in the spatial domain, IEEE Trans. Image Process. 21 (12) (2012) 4695–4708.
[24] G. Yue, C. Hou, K. Gu, N. Ling, No reference image blurriness assessment with local binary patterns, J. Vis. Commun. Image Represent. 49 (2017) 382–391.
[25] K. Gu, D. Tao, J.-F. Qiao, W. Lin, Learning a no-reference quality assessment model of enhanced images with big data, IEEE Trans. Neural Netw. Learn. Syst. (2017), doi:10.1109/TNNLS.2017.2649101.
[26] G. Yue, C. Hou, K. Gu, S. Mao, W. Zhang, Biologically inspired blind quality assessment of tone-mapped images, IEEE Trans. Ind. Electron. 65 (3) (2018) 2525–2536.
[27] Q. Jiang, F. Shao, W. Lin, K. Gu, G. Jiang, H. Sun, Optimizing multi-stage discriminative dictionaries for blind image quality assessment, IEEE Trans. Multim. (2017), doi:10.1109/TMM.2017.2763321.
[28] J. Wang, A. Rehman, K. Zeng, S. Wang, Z. Wang, Quality prediction of asymmetrically distorted stereoscopic 3D images, IEEE Trans. Image Process. 24 (11) (2015) 3400–3414.
[29] F. Shao, K. Li, W. Lin, G. Jiang, M. Yu, Q. Dai, Full-reference quality assessment of stereoscopic images by learning binocular receptive field properties, IEEE Trans. Image Process. 24 (10) (2015) 2971–2983.
[30] F. Shao, W. Lin, S. Wang, G. Jiang, M. Yu, Blind image quality assessment for stereoscopic images using binocular guided quality lookup and visual codebook, IEEE Trans. Broadcast. 61 (2) (2015) 154–165.
[31] F. Qi, D. Zhao, W. Gao, Reduced reference stereoscopic image quality assessment based on binocular perceptual information, IEEE Trans Multim. 17 (12) (2015) 2338–2344.
[32] Z. Wan, F. Qi, Y. Liu, D. Zhao, Reduced reference stereoscopic image quality assessment based on entropy of classified primitives, in: Multimedia and Expo (ICME), 2017 IEEE International Conference on, IEEE, 2017, pp. 73–78.
[33] Q. Jiang, F. Shao, W. Lin, G. Jiang, Learning a referenceless stereopair quality engine with deep nonnegativity constrained sparse autoencoder, Pattern Recognit. 76 (2018) 242–255.
[34] L. Liu, B. Liu, C.-C. Su, H. Huang, A.C. Bovik, Binocular spatial activity and reverse saliency driven no-reference stereopair quality assessment, Signal Process. Image Commun. 58 (2017) 287–299.
[35] W. Zhou, L. Yu, Binocular responses for no-reference 3D image quality assessment, IEEE Trans Multim. 18 (6) (2016) 1077–1084.
[36] W. Zhou, W. Qiu, M.W. Wu, Utilizing dictionary learning and machine learning for blind quality assessment of 3-D images, IEEE Trans. Broadcast. 63 (2) (2017) 404–415.
[37] J. Wang, S. Wang, Z. Wang, Asymmetrically compressed stereoscopic 3D videos: quality assessment and rate-distortion performance evaluation, IEEE Trans. Image Process. 26 (3) (2017) 1330–1343.
[38] L. Shen, J. Lei, C. Hou, No-reference stereoscopic 3D image quality assessment via combined model, Multimed. Tools Appl. (2017), doi:10.1007/s11042-017-4709-7.
[39] F. Shao, Z. Zhang, Q. Jiang, W. Lin, G. Jiang, Towards domain transfer for no-reference quality prediction of asymmetrically distorted stereoscopic images, IEEE Trans. Circuits Syst. Video Technol. 28 (3) (2018) 573–585.
[40] Y. Liu, L.K. Cormack, A.C. Bovik, Statistical modeling of 3-D natural scenes with application to Bayesian stereopsis, IEEE Trans. Image Process. 20 (9) (2011) 2515–2530.
[41] S.K. Md, B. Appina, S.S. Channappayya, Full-reference stereo image quality assessment using natural stereo scene statistics, IEEE Signal Process. Lett. 22 (11) (2015) 1985–1989.
[42] D.L. Ruderman, The statistics of natural images, Network 5 (4) (1994) 517–548.
[43] J. Ding, S.A. Klein, D.M. Levi, Binocular combination of phase and contrast explained by a gain-control and gain-enhancement model, J. Vis. 13 (2) (2013). 13–13.
[44] C.-C. Su, A.C. Bovik, L.K. Cormack, Natural scene statistics of color and range, in: Image Processing (ICIP), 2011 18th IEEE International Conference on, IEEE, 2011, pp. 257–260.
[45] D.J. Fleet, H. Wagner, D.J. Heeger, Neural encoding of binocular disparity: energy models, position shifts and phase shifts, Vis. Res. 36 (12) (1996) 1839–1857.
[46] F. Shao, W. Lin, S. Gu, G. Jiang, T. Srikanthan, Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics, IEEE Trans. Image Process. 22 (5) (2013) 1940–1953.
[47] D. Marr, Vision, San Francisco, CA, USA:Freeman, 1980.
[48] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Trans. Syst. Man Cybern. smc-3 (6) (1973) 610–621.
[49] Z. Wang, E.P. Simoncelli, A.C. Bovik, Multiscale structural similarity for image quality assessment, in: Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on, 2, IEEE, 2003, pp. 1398–1402.
[50] C.-C. Chang, C.-J. Lin, LIBSVM: a library for support vector machines, ACM Trans. Intell. Syst. Technol. 2 (3) (2011) 27.
[51] A.K. Moorthy, C.-C. Su, A. Mittal, A.C. Bovik, Subjective evaluation of stereoscopic image quality, Signal Process. Image Commun. 28 (8) (2013) 870–883.
[52] A. Benoit, P. Le Callet, P. Campisi, R. Cousseau, Quality assessment of stereoscopic images, EURASIP J. Image Video Process. 2008 (1) (2009) 659024.

[53] A. Mittal, R. Soundararajan, A.C. Bovik, Making a "completely blind" image quality analyzer, IEEE Signal Process. Lett. 20 (3) (2013) 209–212.
[54] B. Appina, S. Khan, S.S. Channappayya, No-reference stereoscopic image quality assessment using natural scene statistics, Signal Process. Image Commun. 43 (2016) 1–14.
[55] F. Shao, W. Lin, S. Wang, G. Jiang, M. Yu, Q. Dai, Learning receptive fields and quality lookups for blind quality assessment of stereoscopic images, IEEE Trans. Cybern. 46 (3) (2016) 730–743.
[56] M.-J. Chen, L.K. Cormack, A.C. Bovik, No-reference quality assessment of natural stereopairs, IEEE Trans. Image Process. 22 (9) (2013) 3379–3391.