CrossMark

Quality assessment for virtual reality technology based on real scene

Bin Jiang[1] · Jiachen Yang[1] · Na Jiang[2] · Zhihan Lv[3] · Qinggang Meng[4]

Abstract Virtual reality technology is a new display technology, which provides users with real viewing experience. As known, most of the virtual reality display through stereoscopic images. However, image quality will be influenced by the collection, storage and transmission process. If the stereoscopic image quality in the virtual reality technology is seriously damaged, the user will feel uncomfortable, and this can even cause healthy problems. In this paper, we establish a set of accurate and effective evaluations for the virtual reality. In the preprocessing, we segment the original reference and distorted image to binocular regions and monocular regions. Then, the Information-weighted SSIM (IW-SSIM) or Information-weighted PSNR (IW-PSNR) values over the monocular regions are applied to obtain the IW-score. At the same time, the Stereo-weighted-SSIM (SW-SSIM) or Stereo-weighted-PSNR (SW-PSNR) can be used to calculate the SW-score. Finally, we pool the stereoscopic images score by combing the IW-score and SW-score. Experiments show that our method is very consistent with human subjective judgment standard in the evaluation of virtual reality technology.

1 Introduction

Virtual reality (VR) technology is a way to create and experience the virtual world. Computer system generates a simulation environment which is a fusion of multi-source information. Users can experience the interactive operation just like in the real world [1, 2].

Virtual reality technology mainly includes simulation environment, perception, natural skills and sensor design. The simulation environment is produced by real-time dynamic 3D stereo images [3]. In addition to the computer graphics technology generated by the visual perception, there are hearing, touch, force perception, motion and other sensing [4]. The sense of smell and taste can also be produced. Natural skill refers to the rotation of the head, eyes, gestures or other human behavior action. The user's input makes real-time response [5]. Further more, the sensing device is a three-dimensional interactive device, which is the content we will discuss in this paper [3].

The most acceptable information in a human sense organ is visual through our eyes [6, 7]. Both in the production and life, people need more and more rich visual information to enrich our world and life [4, 7]. As known, the real world is 3D three-dimensional scene. However, the majority of existing 3D stereo display device can only display 2D information, which cannot give people a deep feeling [4]. In order to show the scene and the object with a deep feeling, people attempt to research on all aspects. 3D stereo display technology research has experienced more than ten years of development, and fruitful achievements

✉ Zhihan Lv
Z.Lv@cs.ucl.ac.uk

1   School of Electronic Information Engineering, Tianjin University, Tianjin, People's Republic of China

2   Management Services Company of the First Mining Area, Da Gang Oilfield, PetroChina, Tianjin, People's Republic of China

3   Department of Computer Science, University College London, London, UK

4   Department of Computer Science, School of Science at Loughborough University, Loughborough, UK

have been made. 3D stereo glasses and head display equipment are used for 3D stereo display.

For the display of the world development, 3D has become a well deserved key word. TV display and the other display devices have all launched the 3D stereo display device. Further more, virtual reality has become the next focus of the display device. Although many people have access to 3D information, most of them do not know the specific process. In order to answer these questions, a comprehensive understanding system is needed is shown in Fig. 1.

As known, most of the virtual reality display [5] through the way of stereo images, which is shown in Fig. 1. However, the collection, storage and transmission process of stereoscopic images quality will be influenced in a certain degree [8–10]. If the stereoscopic image quality in virtual reality is seriously damaged, the user will feel uncomfortable, and even cause health problems. So the stereoscopic image quality is the key index to evaluate the quality of a virtual reality system.

Quality assessment of stereoscopic image can be divided into subjective and objective quality evaluations [11, 12]. Subjective quality evaluation method allow observers to give quality score of test according to subjective perception [13, 14]. Although its score is closer to the true visual perception, it wastes time and human resources. Objective quality evaluation method constructs a model that calculates image quality by means of mathematical formula [15–19]. Its advantage is the fast speed, low cost, easy to implement, so it is more practical [20, 21]. At present, there are a lot of researches on the objective evaluation of stereo image quality [22–24]. Some papers construct the models of human visual system (HVS), which can evaluate the quality of the stereoscopic image [25]. For example, the work [26, 27] can evaluate the quality of the stereoscopic image by the combination of the absolute difference of the left and right views on the stereoscopic perception; the

other studies use error calculation (PSNR) as a index to evaluate the quality of the stereoscopic image by simulate multi-channel decomposition characteristics [28, 29]. However, at present, the human understanding of HVS is very limited, so some studies use a simple square error algorithm mean, such as the use of the literature in [30] using PSNR method to define index of Stereo stereoscopic quality objective evaluation from image quality and sense of three-dimensional. In addition, SSIM (Structural Similarity) algorithm [31, 32] is simple and efficient, and the performance is better than PSNR and some evaluation methods such as HVS. So some studies use the classic SSIM algorithm [33–35].

The rest of this paper is organized as follows. Section 2 describes the background and motivation for the idea. Section 3 makes a proposed stereoscopic image quality assessment algorithm in detail. Section 4 shows the experimental results. Section 5 concludes the paper.

2 Background and motivation

2.1 Content-weighted

Most of the image quality assessment algorithms focus on the overall information of the image, while ignoring the difference of local information. In particular, in the process of stereo image processing and evaluation, local high-frequency information may have a great impact on the overall sense of stereo and comfort. Since there is less research in this area, we focus on this direction. The saliency region of the image is the main signal of human visual reception, for example, the strong edge occupies a large proportion of the image information, and it is very important in the visual area [30, 36–39].

In the existing research work, there are many methods about pooling, which can be divided into Minkowski

Fig. 1 The position of stereoscopic image quality in virtual reality, the *red part* is directly related to the content of this paper (color figure online)
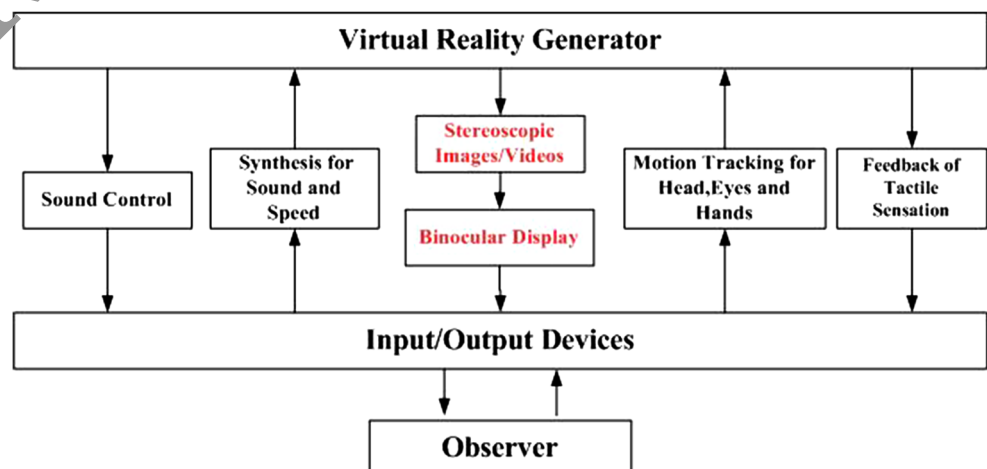
pooling, Local distortion-based pooling, Saliency-based pooling, Object-based pooling and so on [15, 40].

These methods have achieved certain results. However, due to the sensitivity of the database, the pooling method has a dependence on the computational model. And the content of the image is not taken into account.

As a very famous theory, the problem can be well explained by GSM model [15], as described in Eq. 1,

$$R = sU \qquad (1)$$

Among them, R can represent the transform domain vector group of the reference image, U is a zero-mean Gaussian vector, and s can be regarded as a mixing multiplier. Similiarly, it can also be used to represent the distorted image in Eq. 2,

$$D = gR + V \qquad (2)$$

where g is a gain factor and V is additive Gaussian noise. After that, the reference and distorted images pass through a process, and we can know the perceive signals in Eqs. 3 and 4.

$$E = sU + N_1 \qquad (3)$$

$$F = gsU + V + N_2 \qquad (4)$$

where N_1 and N_2 are the Gaussian noise just like V. Based on the theory in [13], we can obtain the mutual information $I(R; E)$ and $I(D, F)$. At last, we can know the information weighted map through Eq. 5.

$$\omega = I(R, E) + I(D, F) - I(E, F) \qquad ( )$$

In the content-weighted IQA, we solve the evaluation problem from image detail information, as shown in Fig. 2. Just like the 2D image evaluation, it is generally considered that visual system Human are optimized information feature extraction process. When using computational simulation of this process, we assume the region which contains more information in the image will

cause much more attention from the observer. Specific to the pixel level, local information content can be used as the weight of the image quality assessment. In stereo image evaluation, the contribution of different regions to the three-dimensional degree will be different, which will be discussed in Sect. 3. So it is necessary to divide the stereo image into binocular correlation region and monocular region.

2.2 Structural similarity

This is a usual method used to evaluate the quality of the 2D image. Because it is very easy for the human vision to draw out the structure information from the image, the similarity of the structure information between the two images can be used to detect the image quality [31].

First structural information should not be influenced by the lighting effects, so it is needed to remove the brightness information in the calculation of structural information. Secondly, structural information should be subject to the influence of image contrast, and the computational structure information needs the difference of the normalized images. Finally we can simply calculate the correlation coefficient of the two pictures after processing the image for the image structure information [32].

However, the image quality is the restriction of brightness and contrast information. So in the calculation of the image quality, information structure also needs to consider the influence. We can commonly use calculation method as Eq. 6,

$$\mathrm{SSIM}(x, y) = [l(x, y)]^{\alpha}[c(x, y)]^{\beta}[s(x, y)]^{\gamma} \qquad (6)$$

These three parts can be calculated separately in Eqs. 7, 8 and 9,

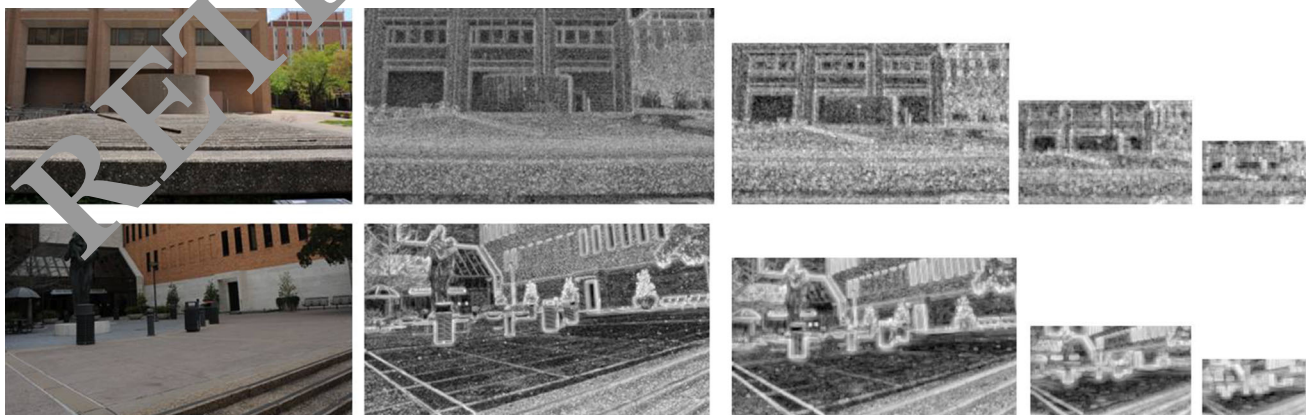$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \qquad (7)$$

Fig. 2 The content-weighted images for different scales, we select two left image from the LIVEI stereoscopic images databases randomly

$$c(x, y) = \frac{2\,\delta_x\,\delta_y + C_2}{\delta_x{}^2 + \delta_y{}^2 + C_2} \tag{8}$$

$$s(x, y) = \frac{\delta_{xy} + C_3}{\delta_x\,\delta_y + C_3} \tag{9}$$

where $l(x, y)$, $c(x, y)$, $s(x, y)$, respectively, the image brightness function, contrast function and structural similarity function of the reference images and the distortion images. C_1, C_2 and C_3 are very small, mainly used to prevent the denominator is zero and instability caused by the phenomenon, $C_1 = (k_1 L)^2$, $C_2 = (k_2 L)^2$, $C_3 = (C_2/2)^2$, k_1, k_2 are far less than the normal number of 1, L is the value of the pixel on the way the range. α, β and γ are parameters to control the weight of three functions, which are greater than 0, generally, $\alpha = \beta = \gamma = 1$. μ_x, μ_y are the mean of x and y, δ_x, δ_y are the variance of x and y, δ_{xy} is the covariance of x and y. We can caculate them in Eqs. 10, 11 and 12.

$$\mu_x = \sum_{i=1}^{N} \omega_i x_i \tag{10}$$

$$\delta_x = \left(\sum_{i=1}^{N} \omega_i (x_i - \mu_i)^2 \right)^{1/2} \tag{11}$$

$$\delta_{xy} = \sum_{i=1}^{N} \omega_i (x_i - \mu_x)(y_i - \mu_x) \tag{12}$$

In the end, we can get the expression of SSIM in Eq. 13.

$$SSIM(x, y) = \frac{1}{M} \sum_{i=1}^{M} SSIM(x_i, y_i). \tag{13}$$

2.3 Binocular vision

Binocular stereo vision is an important form of machine vision, and it is based on the principle of combination for two images from different eyes. Through the position deviation of calculated image corresponding points, we can obtain geometrical information of 3D objects [23]. Image fusion of two eyes is obtained and the difference between them is observed, so that we can get a sense of the apparent depth, which establish the correspondence between features. The distance difference between the same pixel in the left and right views shows the difference between the two eyes. And we call the difference parallax (disparity) image [22].

Binocular vision contains a large amount of information, so it is necessary to map the common information in an image which can reflect left image and right image. The appearance of the cyclopean map has solved this problem [24].

The pioneering work of binocular stereo vision began by Roberts at the Massachusetts Institute of Technology.

Masking of distorted stereoscopic images and binocular rivalry is important in binocular vision. So simulating the true cyclopean image reflects it perfectly by accounting for the display geometry, the presumed fixation, vergence and accommodation.

Firstly, Gabor filter can be used as a simulation of binocular vision, as shown in Eq. 14.

$$G(x, y, \sigma_x, \sigma_y, \xi_x, \xi_y, \theta)$$
$$= \frac{1}{2\pi\sigma_x\sigma_y} e^{-(1/2)[(R_1/\sigma_x)^2 + (R_2/\sigma_y)^2]} e^{i(x\xi_x + x\xi_y)} \tag{14}$$

Then, we can get the cyclopean map by Eqs. 15 and 16, as shown in Fig. 3.

$$\omega_l E_l + \omega_r E_r = C \tag{15}$$

$$Cl(x, y) = W_L(x, y) \times I_L(x, y) + W_R(x + d, y) \times I_R(x + d, y) \tag{16}$$

In the process of stereo imaging, the deviation from the binocular angle will bring the occlusion. The pixel information in a visual angle may not be reflected in another perspective. Then, the process of estimating the disparity map, the pixels are considered to be irrelevant. Mathematical language to express it can be written in Eq. 17,

$$NC_r(x, y) = \begin{cases} 1, & \text{occluded} \\ 0, & \text{else} \end{cases} \tag{17}$$

Similarly, if two points of view can match the region, we think it is relevant in Eq. 18,

$$CR_e(x, y) = \begin{cases} 1, & \text{matched} \\ 0, & \text{else} \end{cases} \tag{18}$$

3 Proposed stereoscopic image quality assessment metric

In this paper, we establish a set of accurate and effective evaluations for the virtual reality, as shown in Fig. 4. First, we divide the original reference and distorted image into two regions called binocular regions and monocular regions. Then, the IW-SSIM or IW-PSNR values over the two regions are applied to certain weights. Finally, we are pooling the stereoscopic images score by combing the two regions.

3.1 Classification for binocular and monocular components

In other people's research work, most of the disparity map is estimated using stereo matching algorithm, and the whole image is decomposed in the process. So, our task is

Fig. 3 The disparity and cyclopean images for stereoscopic images, the *first row* is left images from stereoscopic pairs, the *second row* is the disparity maps corresponding to the *first row*, and the *last row* are cyclopean maps

to detect which areas of the image are correlated with the strong three-dimensional sense.

According to the second part of the paper, for stereo image quality evaluation, binocular region determines the three-dimensional sense, and the other part determines the quality of the 2D images. So we use the Hough transform to detect the edge information of the stereo image. On the basis of this, the classification for binocular region is carried out [41, 42].

Hough transform is a feature extraction technology in image processing, and it detects the object with a specific shape by a voting algorithm. Hough transform is put forward for the first time by Paul Hough. Promotion was made by Richard Duda and Peter Hart. The classical Hough transform for line detection in images and then Hough transform are extended to the recognition of objects of arbitrary shape, circle and ellipse, which can be used in this paper.

We know that a straight line in the Cartesian coordinate system can use $y = kx + b$, the main idea of Hough transform is exchanging the equation parameters and variables. x, y are changed as known, and amount of k, b are changed as the coordinate variables. So the line $y = kx + b$ in Cartesian coordinates can be changed, and a (x_1, y_1) point in Cartesian coordinates is put into a straight line $y_1 = kx_1 + b$. The (k, b) is an arbitrary point on the straight line. In order to facilitate the calculation, we will coordinate the parameter space representation for polar coordinates γ and θ. Because the line with a corresponding points (γ, θ) is the same, it can finish the image edge detection.

Using this method, we can easily divide the three-dimensional image. For different regions, we will use different calculation methods to give the weight of different sizes, as shown in Fig. 5.

3.2 Content-weighted SSIM for monocular component

Because the stereoscopic image is made up of the left and right visual angles, it is not complete to evaluate the visual information of a visual angle separately. However, many methods of plane image evaluation can still be used for reference. In this part, we consider the left and right angles of view as a plane image, evaluate the weight of the information of the plane image and obtain the overall evaluation results. This process fully reflects the three-dimensional image of the inside and the three-dimensional degree of irrelevant content.

By calculating, we can clearly see that the two methods can independently represent the quality of the image, as described in the previous section. So the validity of this

SW-SSIM

Left Referenced

Left Distorted

Reference Cyclopean

Left-IW-SSIM

Distorted Cyclopean

Right Referenced

Right Distorted

Right-IW-SSIM
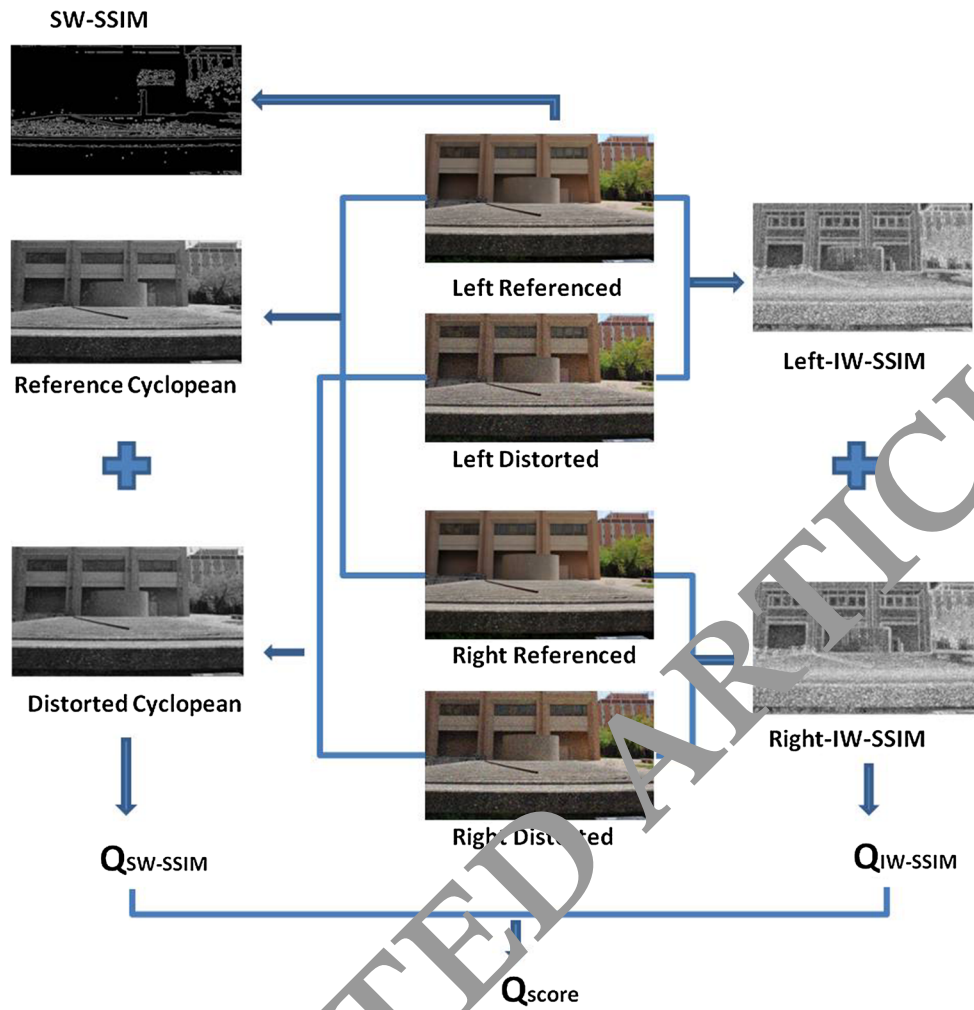
$Q_{SW-SSIM}$

$Q_{IW-SSIM}$

Q_{score}

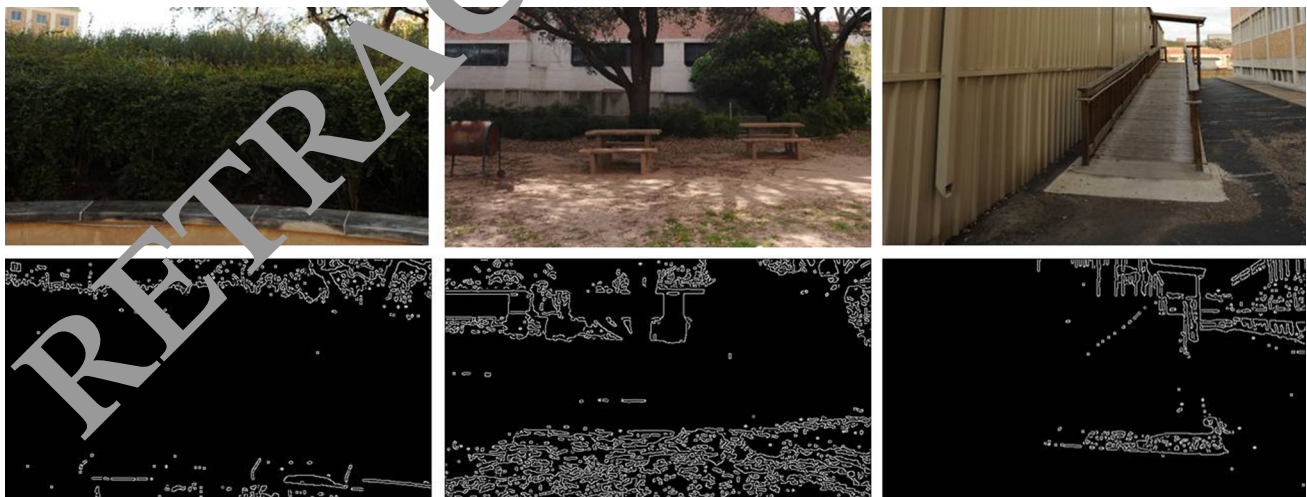**Fig. 4** Framework of the proposed virtual reality image quality assessment metric

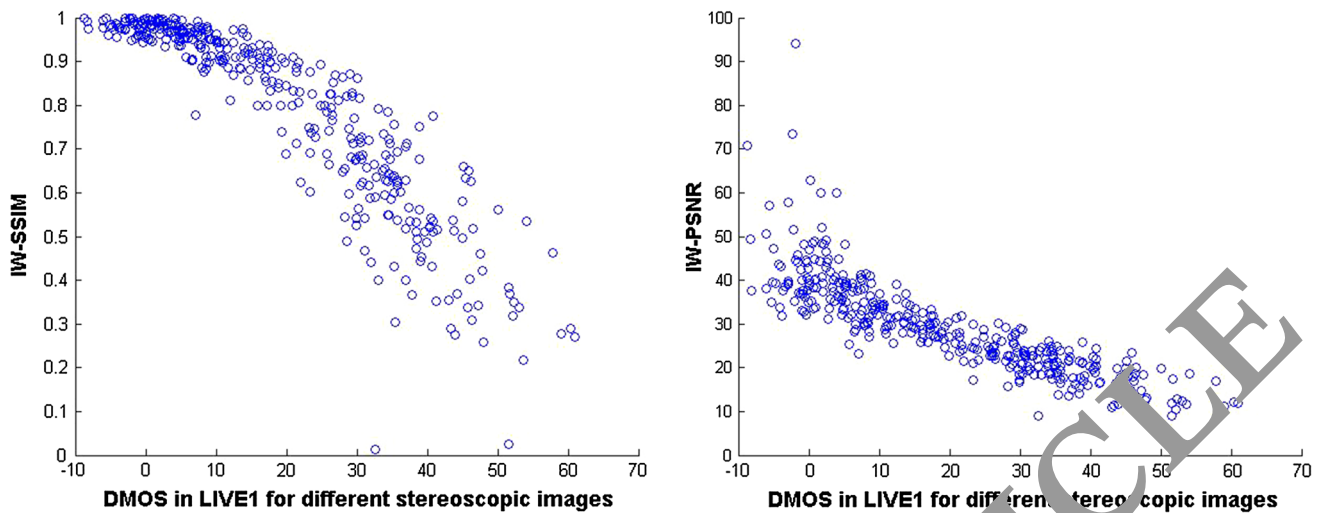**Fig. 5** Detection results of binocular information region of stereo images using Hough transform

Fig. 6 The IW-SSIM and IW-PSNR score for stereoscopic images when we caculate *left* and *right* images as dependent parts

method has been fully demonstrated when we evaluate the stereoscopic image to be divided into two parts in Eq. 19.

$$Q_{\text{IWSSIM}} = \omega_1 \cdot \text{IWSSIM}_{\text{left}} + \omega_2 \cdot \text{IWSSIM}_{\text{right}} \qquad (19)$$

In this part, we take $\omega_1 = \omega_2 = 0.5$. The IW-SSIM and IW-PSNR scores for stereoscopic iamges can be compared in Fig. 6.

3.3 Stereo-weighted SSIM for cyclopean considering binocular vision

As the content of the Hough transform can fully reflect the high-frequency information in the stereoscopic image, it is proved by experiments that the information is very large. So in the use of Hough transform after the calibration area, which needs to be given a greater weight, we can describe it as Eq. 20.

The cyclopean maps contain the left and right image mutual information, which for three-dimensional evaluation is vital, so we also introduce cyclopean images, as a basis for further evaluation [5].

$$Q_{\text{SWSSIM}} = \omega(x,y)\text{IWSSIM}(\text{CYC}_{\text{reference}}, \text{CYC}_{\text{distortion}})$$
$$(20)$$

3.4 Final predict

Finally, $Q_{\text{IW-SSIM}}$ and $Q_{\text{SW-SSIM}}$ are combined into an overall quality score Q in Eq. 21.

$$Q = \omega \cdot Q_{\text{IWSSIM}} + (1 - \omega)Q_{\text{SWSSIM}} \qquad (21)$$

After the weight search calculation, we get the value of ω that is 0.6527. In the next part of the test, we will also prove the rationality of this parameter.

4 Experimental results and analyses

4.1 Stereo image database

In this paper, the process of establishment uses the three-dimensional image database of University of Texas-Austin called 3D Image LIVE database [43], which can be divide into LIVEI and LIVEII. Quality validation of the proposed model can be convinced.

Stereo image database LIVEI contains 20 scenes from the different scenes. LIVEI contains 365 stereoscopic image pairs, each image has its own subjective evaluation, and the distortion of the image is symmetrical. A total of 5 types of distortion are included: JPEG, JP2K, WN, FF and Blur. The stereoscopic image pairs with different types of distortion can be used to defect the performance of detection algorithms in different situations [43].

Stereoscopic image database LIVEII contains 5 types of distortion types just like LIVEI, a total of 360 distortion stereo image pairs, each image corresponding to their respective subjective assessment value.

4.2 Performance measures

In this paper, four kinds of commonly used evaluation indexes are evaluated on the basis of two kinds of stereoscopic image database. The overall performance of the proposed model, that is, Pearson linear correlation coefficient (PLCC), Spearman rank-order correlation coefficient (SROCC), Kendall rank-order correlation coefficient (KROCC) and Root mean square error (RMSE). The best performance corresponds to the PLCC = SROCC = KROCC = 1 and RMSE = 0.

At the same time, the 5 parameters [44] of the linear regression function are used in this paper, as Eq. 22

Table 1 Performance in LIVEI

	PLCC	SROCC	KROCC	RMSE
FI-SSIM	0.8699	0.8606	0.6659	8.0874
FI-VIF	0.9222	0.9188	0.7429	6.3423
FI-UQI	0.8996	0.9022	0.7207	7.1619
Benoit	0.8836	0.8854	0.6876	7.6784
You	0.9312	**0.9248**	0.7503	5.9774
Bensalma	0.8887	0.8767	0.6780	7.5194
Chen	0.9181	0.9090	0.7256	6.5008
Proposed	**0.9322**	0.9206	**0.7831**	**5.8387**

Bold values represent the top performed metric

Table 2 Performance in LIVEII

	PLCC	SROCC	KROCC	RMSE
FI-SSIM	0.6844	0.6795	0.5042	8.2295
FI-VIF	0.7234	0.7213	0.5525	7.7936
FI-UQI	0.7775	0.7720	0.5880	7.0990
Benoit	0.7642	0.7475	0.5553	7.2806
You	0.7744	0.7206	0.5361	7.1413
Bensalma	0.7699	0.7513	0.5761	7.2035
Chen	0.9065	0.9013	0.7307	**4.7665**
Proposed	**0.9167**	**0.9056**	**0.7741**	6.6437

Bold values represent the top performed metric

$$\mathrm{DMOS}_p = \beta_1 \left[\frac{1}{2} - \frac{1}{1+\exp(\beta_2(x-\beta_3))} \right] + \beta_4 x + \beta_5 \tag{22}$$

where β_1, β_2, β_3, β_4 and β_5, can be determined by using the objective scores and subjective scores, and then, the experimental data can be calculated. The result in this part can be seen in Tables 1 and 2.

4.3 Performance for different distortion types

Theoretically speaking, for different types of distortion, stereoscopic image quality assessment should be able to solve the classification. But in the current research, results are from the overall point of view in the consideration of this issue. Evaluation of the index is also for the entire distortion of the three-dimensional image library. So it is necessary for us to compare different types of distortion. In LIVEI, we compare the results obtained by different algorithms on different distortion types. Since PLCC is the most representative of the parameters in the above mentioned parameters, we are here to serve as the main basis for judging. The results of this part are shown in Table 3.

Table 3 PLCC on each individual distortion type in LIVEI

	JEPG	JP2K	WN	GB	FF
FI-SSIM	0.2741	0.8210	0.9250	0.9080	0.7297
FI-VIF	0.6545	0.9412	0.9310	0.9573	0.7572
FI-UQI	0.7764	**0.9532**	0.9374	**0.9633**	0.7752
Benoit	0.5537	0.8789	0.9357	0.9181	0.7377
You	0.6287	0.9390	0.9344	0.9549	**0.8618**
Bensalma	0.3868	0.8413	0.9157	0.9374	0.7385
Chen	0.5142	0.9006	0.9291	0.9 7	0.7994
Proposed	**0.7856**	0.9322	**0.9407**	0.9453	0. 354

Bold values represent the top performed metric

Table 4 Performance on each individual distortion type in LIVEI

	PLCC	SROCC	KROCC	RMSE
IW-SSIM only	0.910	0.90	**0.7939**	6.7520
SW-SSIM only	0.8 56	0.8864	0.7881	7.6545
Proposed	.9302	.9206	0.7831	**5.8387**

Bold values represent the top performed metric

Table 5 Performance on each individual distortion type in LIVEII

	PLCC	SROCC	KROCC	RMSE
IW-SSIM only	0.8954	0.8861	0.7739	**6.1548**
SW-SSIM only	0.8806	0.8753	**0.7981**	6.9854
Proposed	**0.9167**	**0.9056**	0.7741	6.6437

Bold values represent the top performed metric

4.4 Influence of the IW-SSIM and SW-SSIM

As mentioned above, IW-SSIM is mainly used to evaluate the quality of the monocular part, while SW-SSIM is mainly used to evaluate the quality of the binocular part. So the contribution of the two parts is the content of our research.

For using only IW-SSIM and using only the SW-SSIM section, we give the performance on the LIVEI and LIVEII libraries, respectively, as shown in Tables 4 and 5.

5 Conclusion

Since the three-dimensional images in virtual reality technology have an important position and stereo images quality directly affects the whole system, we make research on stereo image quality assessment in order to give the user a better virtual reality experience.

According to the stereoscopic images in virtual reality, this chapter firstly introduces the binocular stereo vision

RETRACTED ARTICLE

model, including the binocular model of stereo vision and the weight of single-item information. Then, according to the IW-SSIM and SW-SSIM model, the objective evaluation model of stereoscopic image quality is put forward. This stereoscopic image quality assessment method effectively extends the previous stereo vision saliency model and then combines the single-channel stereo vision model to simulate the human eye stereo information fusion and processing process. After that, final quality evaluation model is set up. The above method effectively combines the characteristics of the human eye stereo vision with the weight of information, which not only greatly reduces the computational complexity, but also reduces the overall evaluation time.

The results show that the evaluation method can effectively evaluate the quality of the stereo image, and the correlation between the objective evaluation value and the subjective evaluation value is stronger. In the end, the stereoscopic vision mechanism of the human eye is the most important part for the virtual reality. However, this field has just started. Generally speaking, the eye stereoscopic vision information processing and three-dimensional information sensing mechanism are still in development stage. Multi-channel visual characteristics, contrast sensitivity visual characteristics and color visual characteristics, still need further research and analysis, which is significant in this field of research.

References

1. Stern P (2014) Virtual reality changes neural firing. Science 346(6213):1075–1075
2. Komura T, Lau RW, Lin MC et al (2015) Virtual reality software and technology. IEEE Comput Graph Appl 35(5):20–21
3. Klinghammer M, Schütz I, Blohm G, Fiehler K (2016) Allocentric information is used for memory-guided reaching in depth: A virtual reality study. Vis Res 129:13–24
4. Zyda M (2005) From visual simulation to virtual reality to games. Computer 38(9):25–32
5. Hartanto D, Kampmann IL, Morina N et al (2014) Controlling social stress in virtual reality environments. PLoS ONE 9(3):e92804
6. Sas GD, Zachmann G (1998) Integrating virtual reality for virtual prototyping. In: Computers and information in engineering conference, pp 387–399
7. Donalek C, Djorgovski SG, Cioc A et al (2014) Immersive and collaborative data visualization using virtual reality platforms. In: IEEE international conference on big data. IEEE Computer Society, pp 609–614
8. Lin Y, Yang J, Lv Z et al (2015) A self-assessment stereo capture model applicable to the internet of things. Sensors 15(8):20925–20944
9. Lu F, Zhao Q, Yang G (2014) A no-reference image quality assessment approach based on steerable pyramid decomposition using natural scene statistics. Neural Comput Appl 26(1):77–90
10. Yang J, Xu R, Lv Z, Song H. Analysis of Camera Arrays Applicable to the Internet of Things. Sensors 16(3):421–432
11. Liu Y, Yang J, Meng Q et al (2016) Stereoscopic image quality assessment method based on binocular combination saliency model. Signal Process 125:237–248
12. Yang J, Lin Y, Gao Z, Lv Z, Wei W, et al. (2015) Quality Index for Stereoscopic Images by Separately Evaluating Adding and Subtracting. PLOS ONE 10(12):1–19
13. Damera-Venkata N, Kite TD, Geisler WS et al (2000) Image quality assessment based on a degradation model. IEEE Trans Image Process A Publ IEEE Signal Process Soc 9(4):636–650
14. Wang Z, Bovik AC (2002) A universal image quality index. IEEE Signal Process Lett 9(3):81–84
15. Wang Z, Li Q (2011) Information content weighting for perceptual image quality assessment. IEEE Trans Image Process A Publ IEEE Signal Process Soc 20(5):1185–98
16. Sheikh HR, Bovik AC, Veciana GD (2005) An information fidelity criterion for image quality assessment using natural scene statistics. IEEE Trans Image Process A Publ IEEE Signal Process Soc 14(12):2117–28
17. Sheikh HR, Bovik AC (2006) Image information and visual quality. IEEE Trans Image Process A Publ IEEE Signal Process Soc 15(2):430–444
18. Chandler DM, Hemami SS (2007) VSNR: a wavelet-based visual signal-to-noise ratio for natural images. IEEE Trans Image Process A Publ IEEE Signal Process Soc 16(9):2284–98
19. Zhang L, Zhang D and Mou X (2010) RFSIM: A feature based image quality assessment metric using Riesz transforms. In: IEEE International Conference on Image Processing, pp 321–324
20. Moorthy AK, Bovik AC (2011) Blind image quality assessment: from natural scene statistics to perceptual quality. IEEE Trans Image Process A Publ IEEE Signal Process Soc 20(12):3350–64
21. Saad MA, Bovik AC, Charrier C (2012) Blind image quality assessment: a natural scene statistics approach in the DCT domain. IEEE Trans Image Process A Publ IEEE Signal Process Soc 21(8):3339–52
22. Lin YH, Wu JL (2014) Quality assessment of stereoscopic 3D image compression by binocular integration behaviors. IEEE Trans Image Process 23(4):1527–42
23. Benoit A, Callet PL, Campisi P et al (2008) Using disparity for quality assessment of stereoscopic images. In: 15th IEEE international conference on image processing, 2008. ICIP 2008. IEEE, pp 389–392
24. Chen MJ, Su CC, Kwon DK et al (2013) Full-reference quality assessment of stereopairs accounting for rivalry. Signal Process Image Commun 28(9):1143–1155
25. Kang L, Ye P, Li Y and Doermann D (2014) Convolutional Neural Networks for No-Reference Image Quality Assessment. In: IEEE conference on computer vision and pattern recognition, pp 1733–1740
26. Zhang W, Qu C, Ma L, Guan J, Huang R (2016), Learning structure of stereoscopic image for no-reference quality assessment with convolutional neural network. Pattern Recognit 59:176–187
27. Shao F, Lin W, Jiang G and Dai Q (2016) Models of Monocular and Binocular Visual Perception in Quality Assessment of Stereoscopic Images. IEEE Trans Comput Imaging 2(2):123–135
28. Tang H, Joshi N, Kapoor A (2014) Blind image quality assessment using semi-supervised rectifier networks. In: IEEE conference on computer vision and pattern recognition, pp 2877–2884
29. Xue W, Zhang L and Mou X (2013) Learning without Human Scores for Blind Image Quality Assessment. In: IEEE Conference on Computer Vision and Pattern Recognition, pp 995-1002
30. Larson EC, Chandler DM (2010) Most apparent distortion: full-reference image quality assessment and the role of strategy. J Electron Imaging 19(1):143–153

31. Wang Z, Bovik AC, Sheikh HR et al (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612

32. Wang Z, Simoncelli EP, Bovik AC (2003) Multiscale structural similarity for image quality assessment. In: Conference record of the thirty-seventh asilomar conference on signals, systems and computers, vol 2. pp 1398–1402

33. Ye P, Doermann D (2011) No-reference image quality assessment based on visual codebook. In: IEEE international conference on image processing, pp 3089–3092

34. Doermann D (2012) Unsupervised feature learning framework for no-reference image quality assessment. In: IEEE conference on computer vision and pattern recognition, IEEE Computer Society, pp 1098–1105

35. Li C, Bovik AC, Wu X (2011) Blind image quality assessment using a general regression neural network. IEEE Trans Neural Netw 22(5):793–799

36. Zhang L, Zhang L, Mou X et al (2011) FSIM: a feature similarity index for image quality assessment. IEEE Trans Image Process A Publ IEEE Signal Process Soc 20(8):2378–86

37. Liu A, Lin W, Narwaria M (2012) Image quality assessment based on gradient similarity. IEEE Trans Image Process A Publ IEEE Signal Process Soc 21(4):1500–12

38. Zhang L, Shen Y, Li H (2014) VSI: a visual saliency-induced index for perceptual image quality assessment. IEEE Trans Image Process 23(10):4270–81

39. Mittal A, Moorthy AK, Bovik AC (2012) No-reference image quality assessment in the spatial domain. IEEE Trans Image Process A Publ IEEE Signal Process Soc 21(12):4695–4708

40. Li C, Bovik AC (2010) Content-weighted video quality assessment using a three-component image model. J Electron Imaging 29(1):143–153

41. Yang D, Guo P (2011) Image modeling with combined optimization techniques for image semantic annotation. Neural Comput Appl 20(7):1001–1015

42. Chen L, Li C, Huang T et al. (2014) Memristor crossbar-based unsupervised image learning. Neural Comput Appl 25(2):393–400

43. Moorthy AK, Su CC, Mittal A et al (2013) Subjective evaluation of stereoscopic image quality. Signal Process Image Commun 28(8):870–883

44. Gottschalk PG, Dunn JR (2005) The five-parameter logistic: a characterization and comparison with the four-parameter logistic. Anal Biochem 343(1):54–65