

Research Article

No-Reference Stereoscopic IQA Approach: From Nonlinear Effect to Parallax Compensation

Ke Gu, Guangtao Zhai, Xiaokang Yang, and Wenjun Zhang

Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, No. 800 Dongchuan Road, Minhang District, Shanghai 200240, China

Correspondence should be addressed to Ke Gu, gukesjtuee@gmail.com

Received 20 April 2012; Revised 31 August 2012; Accepted 22 September 2012

Academic Editor: Paul Dan Cristea

Copyright © 2012 Ke Gu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The last decade has seen a booming of the applications of stereoscopic images/videos and the corresponding technologies, such as 3D modeling, reconstruction, and disparity estimation. However, only a very limited number of stereoscopic image quality assessment metrics was proposed through the years. In this paper, we propose a new no-reference stereoscopic image quality assessment algorithm based on the nonlinear additive model, ocular dominance model, and saliency based parallax compensation. Our studies using the Toyama database result in three valuable findings. First, quality of the stereoscopic image has a nonlinear relationship with a direct summation of two monoscopic image qualities. Second, it is a rational assumption that the right-eye response has the higher impact on the stereoscopic image quality, which is based on a sampling survey in the ocular dominance research. Third, the saliency based parallax compensation, resulted from different stereoscopic image contents, is considerably valid to improve the prediction performance of image quality metrics. Experimental results confirm that our proposed stereoscopic image quality assessment paradigm has superior prediction accuracy as compared to state-of-the-art competitors.

1. Introduction

Three-dimensional (3D) imaging has been an extensive research area, the application of which ranges from entertainment, such as videos and games, to specialized domains, such as education and medicine. As more and more image processing operations have been specifically designed for stereoscopic images, the necessity for an effective perceptual stereoscopic image quality assessment (IQA) algorithm is increasing. Following the research of monoscopic image quality metrics, stereoscopic IQA approaches fall into two categories: subjective assessment and objective assessment. Although the subjective assessment method should be the ultimate quality gauge for digital images, it is usually time-consuming, expensive, and not practical for real-time image processing systems. Therefore, an increasing number of objective stereoscopic image quality metrics have been developed. According to the availability of reference images to be compared with during the tests, objective stereoscopic IQA methods can be further classified into three categories. First, the most general approaches are full-reference methods

[1, 2], assuming the reference image is fully known. In many practical applications, however, the reference image is not available, and the second type of methods, namely, no-reference image quality metrics [3, 4] is then desirable. The third type is referred to as reduced-reference IQA algorithm [5], which is applied to the situation where the reference image is only partially available, that is, some extracted features are made available as side information to help estimate quality of the distorted image. This paper concentrates on the no-reference type of IQA approaches.

Many valuable monoscopic no-reference image quality metrics were proposed during the last decade. Wang et al. and Sheikh et al. proposed IQA methods for JPEG and JPEG2000 compressed images in [6, 7], which have obtained high prediction performance. Besides, Blind Image Quality Indices (BIQI) [8] proposed were based on image distortion classification followed by those in [6, 7] and other image quality metrics used for different specific types of distortion. This no-reference IQA approach not only achieves much better results but also opens a major new direction for the current research of IQA algorithms.

Extension from monoscopic image quality metrics to stereoscopic IQA methods is a challenging work, although some inspiring models have been proposed in [1, 2, 4, 5]. In the real world, it is not difficult to find that the human subjective feelings between two-dimensional (2D) and 3D images are immensely distinct on viewing, but they still have a very close contact with each other. This phenomenon is mainly supported by the fact [9] that cells in the retina of each eye individually encode its received information, and then the information from both eyes is merged in lateral geniculate nucleus (LGN) to form the final stereoscopic image in the brain. Therefore, we were enlightened to design a no-reference stereoscopic IQA paradigm based on 2D image quality metrics and the relationship between 2D and 3D image qualities.

So far, few stereoscopic image quality metrics have been studied on the influence of nonlinear additive model between the left and right image qualities on the perceptual quality of the stereoscopic image, but it is noticed in our research that a kind of nonlinear additive effect mainly takes effect between them. Firstly, we found that the linear additive model can constitute a link between the 2D and 3D image qualities. However, it can hardly delight us due to the demonstration in [10] that the amplitude tends to be reduced with anticorrelated stimuli because of the responses of V1 complex cells. Furthermore, inspired by the research of overlapping effects among different categories of contrast in [11], it is observed that overlapping effects exist between the qualities of left and right images on their integration, especially under the condition that the correlation of their qualities is quite weak. Then, by decreasing these overlapping parts, our proposed nonlinear additive model gains an inspiring improvement.

Through further researches, we found that the above-mentioned nonlinear additive effect is probably caused by the discrepancy of qualities between left and right images. This phenomenon can be explained by some researches on ocular dominance [12–17]. Inspired by Ocular Dominance Index (ODI) in [12, 13], this paper proposes an ocular dominance model, which is defined as ODI weighted quality difference between left and right images. It will be demonstrated widely that our ocular dominance model has significant influences on the perceptual quality of stereoscopic images, just like the way the nonlinear additive model performs. Besides, it is pointed out in [14–17] that about two-thirds of the population is right-eye dominant and one-third is left-eye dominant, while neither eye is dominant in a small portion of the population. So, this conclusion enlightened us to make the right-eye response have higher weight on prediction of the final stereoscopic image quality, which defines our ocular dominance weighting model, and it also contributes a certain promotion.

Finally, different degrees of parallaxes resulted from various stereoscopic image contents also highly affect the prediction accuracy of image quality metrics, based on the experimental result in [18] that the low subjective evaluation is caused by high degree of parallax. Thus, except for the reduction of the nonlinear additive effects, the compensation of different degrees of parallaxes is also applied here. In

addition, visual attention (VA) based IQA methods in [19, 20] illuminated us to introduce the applications of double various VA models [21, 22]. As expected, the parallax compensation can be devoted to much higher prediction results for stereoscopic image quality metrics.

Consequently, based on a 2D no-reference image quality metric ([6] is chosen here) and our proposed effective models, we present a novel nonlinear additive model, ocular dominance (weighting) model, and saliency based parallax compensation based distortion metric (NOSPDM) for JPEG compressed stereoscopic images. Our NOSPDM method basically operates in five steps: individual predictions of left and right image qualities, nonlinear additive model based combination of both quality measures, reduction of the ODI weighted discrepancy between left and right image qualities, different weights for both eyes supported by ocular dominance researches, and, finally, saliency based parallax compensation resulted from different stereoscopic image contents.

The remainder of this paper is organized as follows. In Section 2, the phenomenon of nonlinear additive effect is first reviewed, and then the employment of nonlinear additive model is described in detail. Supported by the ocular dominance researches, Section 3 explicitly presents the proposed ocular dominance (weighting) model. The influence of different parallaxes and the corresponding compensation methodology combined with two classical VA models are mainly introduced in Section 4. Section 5 proposes our NOSPDM paradigm. In Section 6, experimental results using the Toyama database [3] are reported and analyzed. Eventually, conclusion is drawn and future work is discussed in Section 7.

2. Nonlinear Additive Model

For quality assessment of monoscopic images, human observers generally score the quality by quantifying the distortion or difference between the reference and distorted images. However, for stereoscopic images, it has been noticed that two images (left and right images) are individually received by different eyes, and then the final stereoscopic image is formed by merging both monoscopic images in LGN [9]. Thus, under the condition that predictions of both 2D image qualities (illustrated in Section 5.1) are available, a reliable combination model becomes the key point.

It would be natural to employ the linear additive model, one of the simplest models for incorporating two parts. However, we are hardly content with its result, as illustrated in Figure 9 and Table 4 (NOSPDM₁). Then, to find the problem of the linear additive model, eight nonexperienced assessors are invited to score fourteen monoscopic images (Figure 2 and all the other corresponding 2D distorted images) with the same 3D image content, and their subjective scores versus mean opinion score (MOS) values of stereoscopic images in the Toyama database [3] are displayed in Figure 1. Through this test, two significant conclusions can be drawn. Firstly, the relationship between 3D image quality and both 2D image qualities is far beyond the linear additive model.

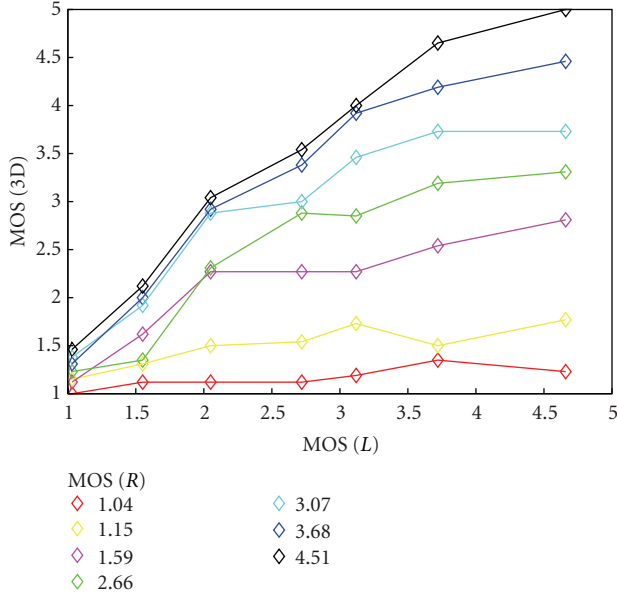


FIGURE 1: Illustration of qualities of stereoscopic images with the same “computer” image content versus qualities of their corresponding left and right images. MOS(3D), MOS(L), and MOS(R) indicate MOS values of stereoscopic, left and right images, respectively.

TABLE 1: Demands for the chosen image groups for testing the “distress” feeling.

Image index	Left image quality	Right image quality
1	Low	High
2	High	Low
3	Fair	Fair

TABLE 2: Demands for the chosen four image pairs with different degrees of JPEG distortion.

Image pairs index	JPEG compressive quality of left image	JPEG compressive quality of right image
1	15	79
2	79	15
3	27	27
4	100	100

Secondly, the stereoscopic image quality is sensitive when the discrepancy between its corresponding two monoscopic image qualities is quite large. For example, the higher value the left image quality becomes, the lower increasing rate of its stereoscopic image quality gets, as marked by black diamond “ \diamond ” in Figure 1. If the different qualities of left and right images can be regarded as different stimuli, the fact stated above basically coincides with the results of experiments in [10], which discovered that amplitude tends to be lessened with anticorrelated stimuli caused by the responses of V1 complex cells.

Further idea is enlightened by [11], in which it is verified that no two saliency effects were found to be strictly independent in all subjects. When viewers score monoscopic images,

TABLE 3: Different combinations of NOSPDM_x. Y represents the corresponding model applied in the (x)th combination.

x -index	F_L	F_O^W	F_N	F_O^D	F_P	F_{VP}	F_{SP}
1	Y						
2		Y					
3	Y		Y				
4		Y	Y				
5	Y			Y			
6		Y		Y			
7	Y		Y		Y		
8	Y		Y			Y	
9	Y		Y				Y
10		Y	Y		Y		
11		Y	Y			Y	
12		Y	Y				Y
13	Y			Y	Y		
14	Y			Y		Y	
15	Y			Y			Y
16		Y		Y	Y		
17		Y		Y		Y	
18		Y		Y			Y

TABLE 4: PLCC, SRCC, KRCC, AAE, and RMSE values (after nonlinear regression) of testing and training groups in [3] and NOSPDM₁–NOSPDM₁₈.

Metrics	PLCC	SRCC	KRCC	AAE	RMSE
[3] (testing)	0.9350	—	—	0.3500	0.4210
[3] (training)	0.9660	—	—	0.2920	0.3670
NOSPDM ₁	0.8934	0.8897	0.7066	0.4136	0.5234
NOSPDM ₂	0.8943	0.8912	0.7097	0.4120	0.5212
NOSPDM ₃	0.9350	0.9312	0.7666	0.3303	0.4133
NOSPDM ₄	0.9360	0.9325	0.7699	0.3269	0.4100
NOSPDM ₅	0.9318	0.9266	0.7609	0.3358	0.4228
NOSPDM ₆	0.9288	0.9231	0.7556	0.3416	0.4316
NOSPDM ₇	0.9493	0.9433	0.7926	0.2851	0.3664
NOSPDM ₈	0.9665	0.9600	0.8284	0.2330	0.2992
NOSPDM ₉	0.9705	0.9652	0.8416	0.2146	0.2807
NOSPDM ₁₀	0.9502	0.9444	0.7951	0.2814	0.3630
NOSPDM ₁₁	0.9673	0.9611	0.8313	0.2289	0.2955
NOSPDM ₁₂	0.9711	0.9660	0.8442	0.2114	0.2782
NOSPDM ₁₃	0.9470	0.9400	0.7878	0.2896	0.3741
NOSPDM ₁₄	0.9638	0.9565	0.8223	0.2402	0.3106
NOSPDM ₁₅	0.9676	0.9613	0.8340	0.2238	0.2942
NOSPDM ₁₆	0.9445	0.9373	0.7837	0.2958	0.3827
NOSPDM ₁₇	0.9610	0.9534	0.8162	0.2476	0.3220
NOSPDM ₁₈	0.9647	0.9581	0.8278	0.2344	0.3067

their saliency regions have highly significant influences on the final subjective scores. Moreover, our used 2D no-reference image quality assessment approach [6] depends on perceiving local distortion, which also belongs to a kind of saliency on a broader definition. Therefore, we have a reason to believe that overlapping effects exist between 3D image

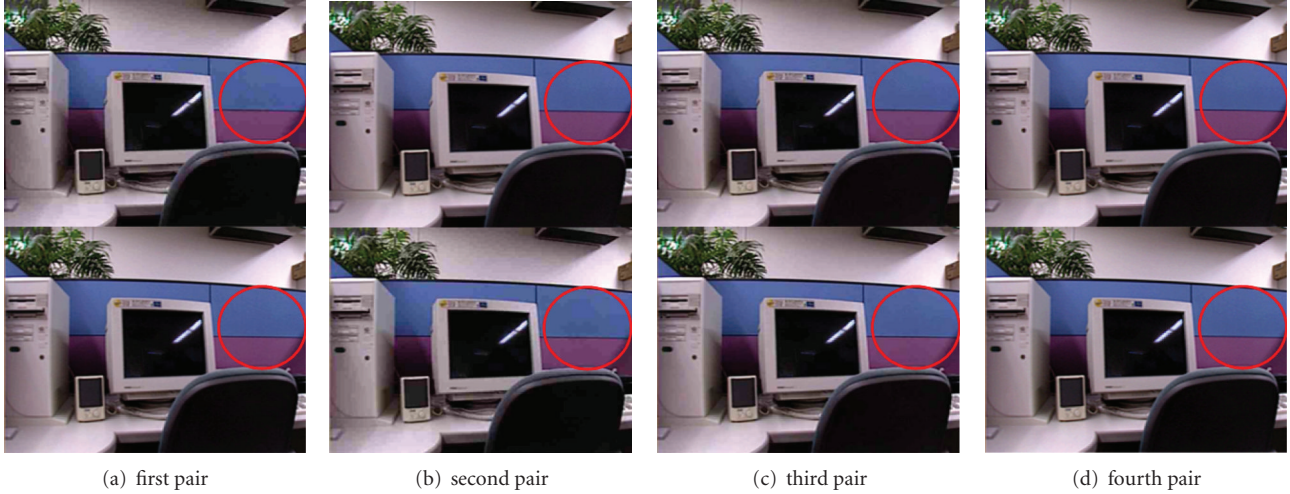


FIGURE 2: Four monoscopic image pairs with the same “computer” image content: ((a)–(c)) three pairs of distorted images; (d) reference image pairs. The regions marked in red circle can clearly show the different JPEG compressive qualities of every image.

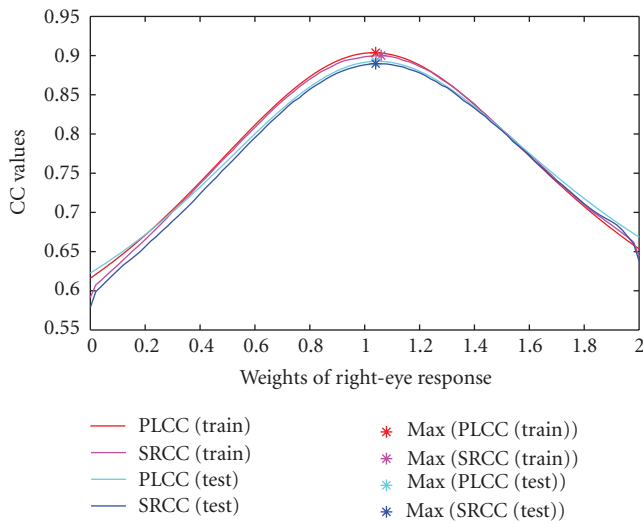


FIGURE 3: PLCC and SRCC values (given in Section 6) based on different weights of right-eye response (the summation of both weights is 2) and their maximum values and corresponding weight values (“*”). Weights of right-eye response are tested from 0 to 2, and the concealed part has the lower value than the revealed part.

quality and both 2D image qualities. Following the method in [23], our nonlinear additive model can be computed as

$$F_N = \lambda \cdot \max\{Q_{2D}(L), Q_{2D}(R)\}, \quad (1)$$

where L and R indicate left and right images, and Q_{2D} represents a 2D image quality assessment method. Here, maximum operator is used to replace minimum operator, since image quality metrics based on measuring difference or distortion mainly pay attention to the lower quality image or the poorer quality regions, which is opposite to the situation of [11, 23].

3. Ocular Dominance (Weighting) Model

It is described in detail above that our proposed nonlinear additive model can overcome the obstacle introduced from the nonlinear additive effect between left and right image qualities. However, this model is only illuminated by some researches about overlapping effects, lacking theoretical principles.

3.1. Compensation of “Distress” of Stereoscopic Image Pairs. Further experiment is to choose more monoscopic image pairs from the Toyama database [3] and divide them into different testing groups. Every group should include three image pairs with the same stereoscopic image content, which meets the requirement presented in Table 1. One exemplary group is shown in Figures 2(a)–2(d).

If we observe the stereoscopic images shown in Figures 2(a)–2(d) with 3D shutter glasses, at a glance, we can immediately tell that the third stereoscopic image has the noticeable distortion. However, we can barely find any difference between the reference and distorted images for the first and second stereoscopic images. This phenomenon can be explained by the fact that there is one image of high quality in the image pairs, and then the visual transience effect makes observers feel good because of the rapid switch of 3D glasses shutters between the open and close states. Contrarily, since the qualities of both images in the third stereoscopic image are not good, they are of much lower perceptual quality.

However, if viewing the first or second stereoscopic image for sufficiently long time, we can feel uncomfortable or even dazzled. The MOS values from the Toyama database verify this observation: generally, the third stereoscopic image quality is the highest of the three, while the MOS value of the second stereoscopic image is a little higher than that of the first one. Demonstrated by many different groups meeting the demand in Table 1, it is noticed that this phenomenon exists widely. Consequently, we believe that the “distress” feeling is caused by the difference between left and

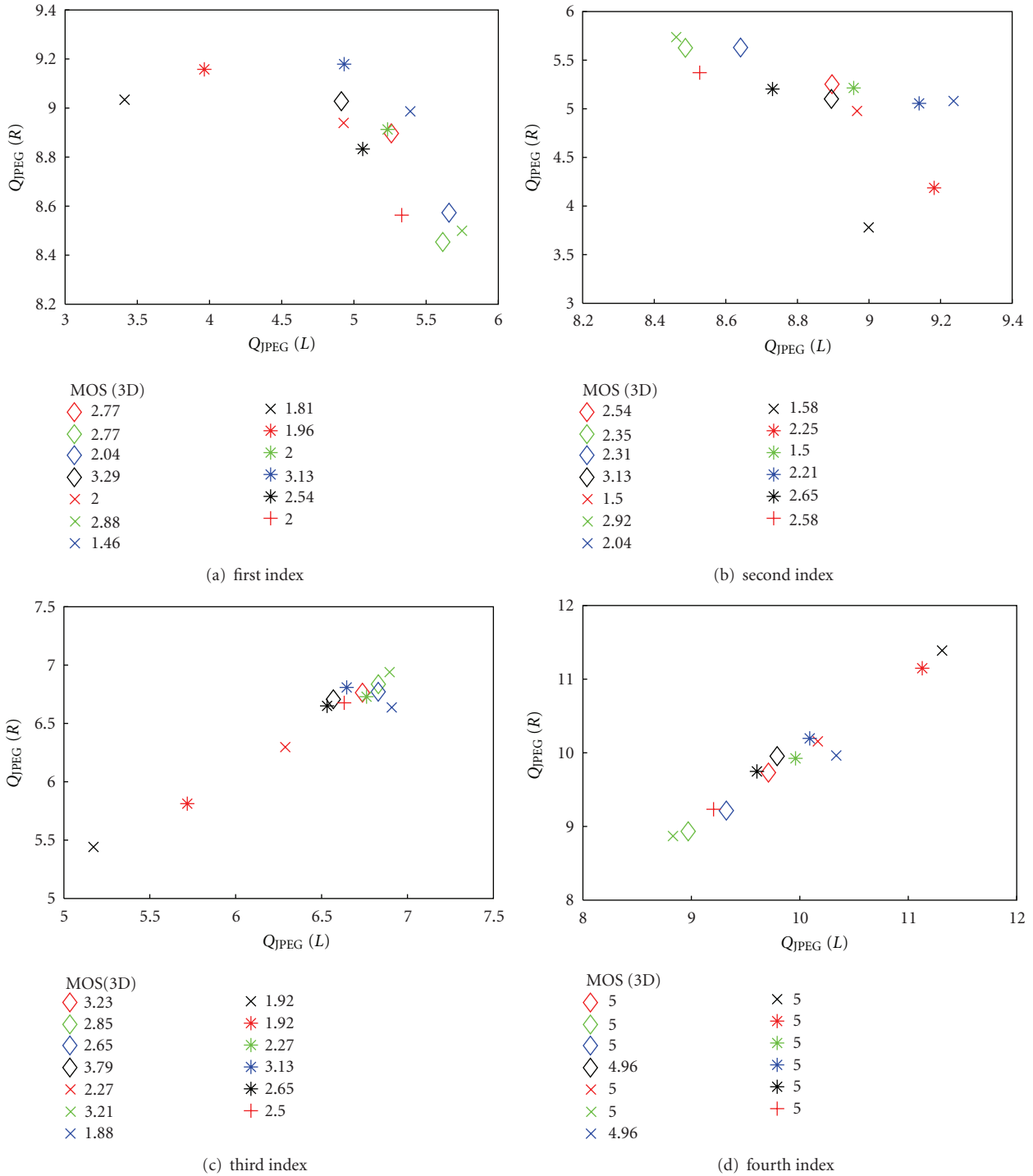


FIGURE 4: Illustration of the (a) first index; (b) second index; (c) third index; (d) fourth index of image pairs in Table 1. $Q_{JPEG}(L)$ and $Q_{JPEG}(R)$ are predictions of left and right image qualities by [6], and MOS(3D) indicates MOS values of stereoscopic images.

right image qualities, and, moreover, the “distress” feeling becomes more serious with their larger discrepancy.

When observing a stereoscopic image, if the left and right images are almost of the same quality, our eyes can function equally, and the images are merged smoothly in LGN [9]. However, when the left and right image qualities are quite

different, for example, the left image is clear but the right one is heavily contaminated as shown in Figures 2(c) and 2(d), the imbalance causes the inequality of the two eyes [24] and LGN will have difficulty in merging left and right images. The brain must tense the muscles around the outside of the eye that is receiving the low quality image so as to

1	2	1
4	8	4
1	2	1

FIGURE 5: A simple visual attention model defined in [21].

steady the view. And this makes the brain suffer. After a while, the muscles start to ache and the nerves begin to pain. And this unbalanced image pair may even result in amblyopia for a very long-term viewing [24].

In this study, this important “distress” is defined as

$$G(X, Y) = H(X, Y) \cdot \|X - Y\|_{\infty}, \quad (2)$$

where X and Y are the left- and right-eye responses respectively, and $H(X, Y)$ denotes a “distress” degree parameter. Furthermore, the interaction between the left- and right-eyes should have impacts on the final perception of the 3D image quality. Inspired by the definition of ODI in [12, 13]

$$\text{ODI} = \frac{[\text{left_eye_response}]}{[\text{left_eye_response}] + [\text{right_eye_response}]}, \quad (3)$$

we believe that ODI is a good characterization of this interaction, and it can constitute a link between both 2D image qualities and the final 3D image quality. So, by taking ODI into $H(X, Y)$ and replacing eye responses X and Y with quality predictions $Q_{2D}(L)$ and $Q_{2D}(R)$ in (2), our ocular dominance model can be given by

$$\begin{aligned} F_O^D &= \text{ODI} \cdot \|X - Y\|_{\infty} = \frac{X}{X + Y} \cdot \|X - Y\|_{\infty} \\ &= \frac{Q_{2D}(L)}{Q_{2D}(L) + Q_{2D}(R)} \cdot \|Q_{2D}(L) - Q_{2D}(R)\|_{\infty}. \end{aligned} \quad (4)$$

3.2. Asymmetric Weights of Different Eye Responses. Besides, an important conclusion is given in [14–17] that about two-thirds of the population are right-eye dominant and one-third is left-eye dominant, while neither eye is dominant in a small portion of the population. Thus, we have a reason to suppose that asymmetric weights of different eye responses have a certain impact on prediction performance of stereoscopic IQA method. This supposition can be testified by using all the stereoscopic images with the same “computer” image content, as partly shown in Figure 2, and their MOS values of the 3D images are shown below:

$$\text{MOS}_{3D} = \begin{bmatrix} 1.00 & 1.12 & \widetilde{1.12} & 1.12 & \overline{1.19} & 1.35 & \overline{1.23} \\ 1.15 & 1.31 & 1.50 & \overline{1.54} & \overline{1.73} & \overline{1.50} & \overline{1.77} \\ \widetilde{1.12} & 1.62 & 2.27 & 2.27 & \overline{2.27} & \overline{2.54} & 2.81 \\ 1.23 & \overline{1.35} & 2.31 & 2.88 & 2.88 & 3.19 & 3.31 \\ \overline{1.38} & \overline{1.92} & \overline{2.88} & 3.00 & 3.00 & 3.73 & 3.73 \\ 1.31 & 2.00 & \overline{2.92} & 3.38 & 3.38 & 4.19 & 4.46 \\ \overline{1.46} & 2.12 & 3.04 & 3.54 & 3.54 & 4.65 & 5.00 \end{bmatrix}. \quad (5)$$

Just as our expectation, it can be seen from the MOS_{3D} matrix above that only one couple of MOS_{3D} values covered

with a wave line (about 5% of all) is consistent with the common sense that two eyes have the same function, 12 couples in bold type (about 57%) is against it (but their deviations are less than 10%), and 8 couples in bold type and covered with a horizontal line (about 38%) show the completely contrary results. Demonstrated by all the other 3D images, this fact is widely existent although some slight differences exist in the percentage distribution of these three categories of results.

To quantify the weights of different eye responses, at first, all the stereoscopic images are randomly divided into two groups (training group and testing group) according to the reference images. Then, the test of finding the most reliable weights is carried out by making NOSPDM_2 (defined by (17) in Section 5.2) achieve the best quality prediction performance in terms of the highest correlation coefficients. Figure 3 displays the above-mentioned results. It can be concluded that the right-eye response should have a weight larger than the left-eye response, and here we set μ_R (the weight of right-eye response) equal to 1.04 (the corresponding weight of maximum values “*” in Figure 3). So, the ocular dominance weighting model can be eventually estimated as

$$\begin{aligned} F_O^W &= \mu_L \cdot Q_{2D}(L) + \mu_R \cdot Q_{2D}(R) \\ &= (2 - \mu_R) \cdot Q_{2D}(L) + \mu_R \cdot Q_{2D}(R), \end{aligned} \quad (6)$$

where μ_L and μ_R represent the weights of left- and right-eye responses.

4. Saliency Based Parallax Compensation

Besides the existing nonlinear additive effects in the monoscopic image pairs with the same stereoscopic image content, which is clarified in Sections 2 and 3, it is also viewed that another kind of additive effect exists among stereoscopic images with different image contents. We first test various groups of 2D image pairs, which cover all the different image contents. Every group of four image pairs should meet the demand illustrated in Table 2, which lists the chosen four categories of image pairs with different degrees of JPEG distortion. And Figure 2 is just one exemplary group.

The experimental results shown in Figure 4 confirm our finding. However, in the meantime, we observed two interesting phenomena. First, against our common sense that the higher qualities of left and right images the higher quality of the corresponding 3D image, the stereoscopic images represented by blue “×” and green “*” in Figure 4(a) with high qualities of both 2D images have the unusually low qualities. The similar phenomenon also appears in Figures 4(b)–4(d). Second, it is also found in Figure 4 that the existence of the above-mentioned phenomenon tends to appear for testing stereoscopic images with the same JPEG compressive quality but different image contents. In the whole Toyama database, the two above-mentioned findings exist to a large extent.

This fact cannot be simply explained by the interaction between qualities of left and right images, otherwise it should not appear in Figure 4(c) (or Figure 4(d)), where each

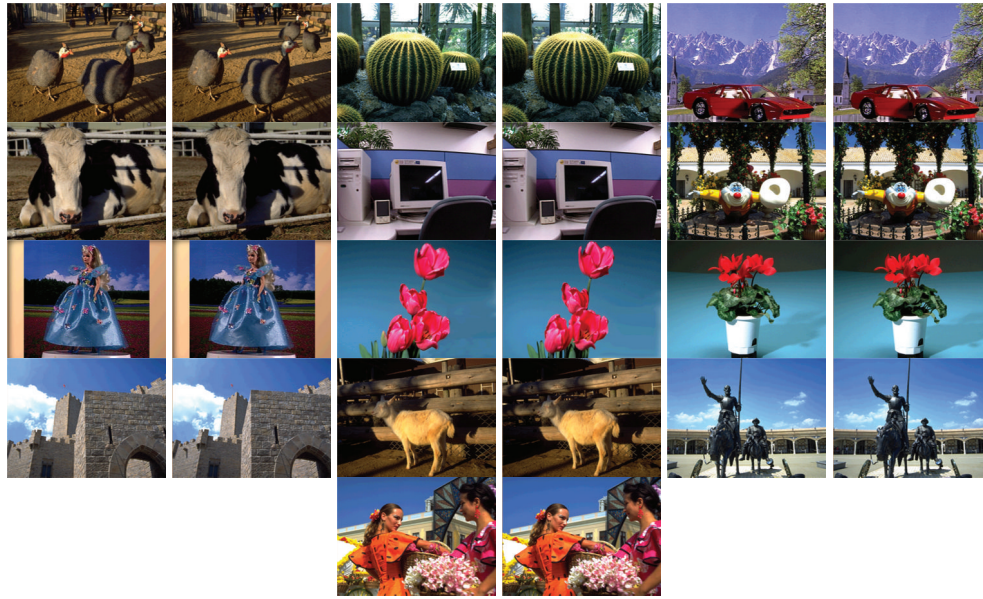


FIGURE 6: Reference stereoscopic images in the Toyama database [3]: bird, cactus, car, cattle, computer, doll, doll2, flower, flower2, gate, goat, statue, and women in turn.

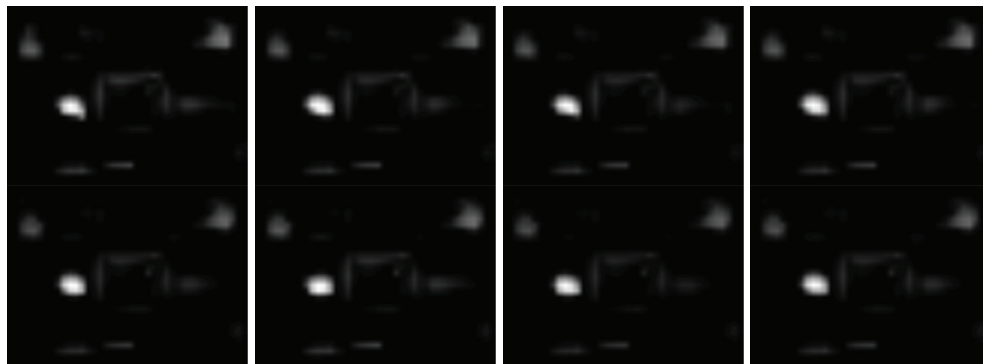


FIGURE 7: Corresponding saliency maps of Figure 2 computed by the authors in [22].

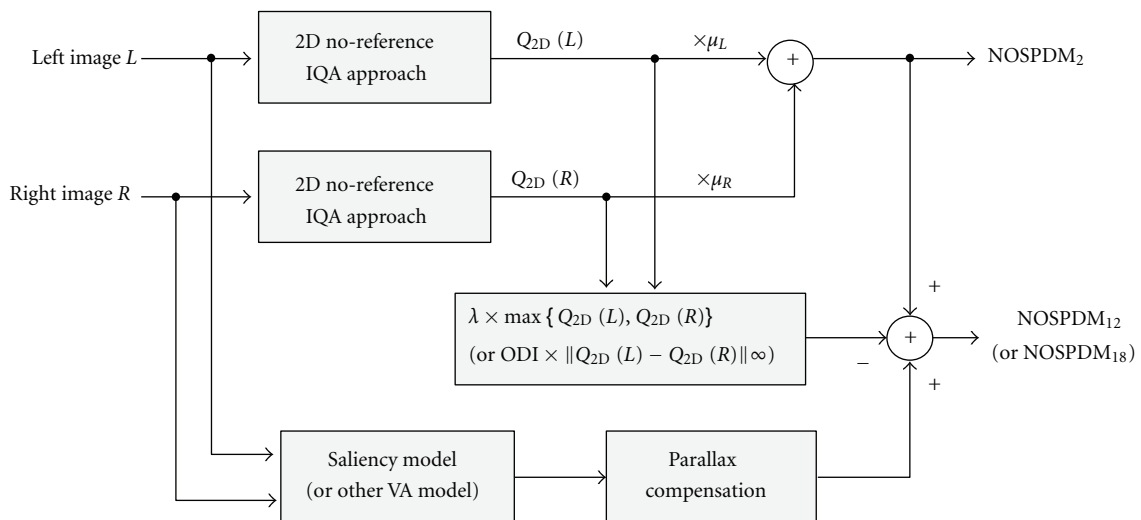


FIGURE 8: Five steps of NOSPDM framework.

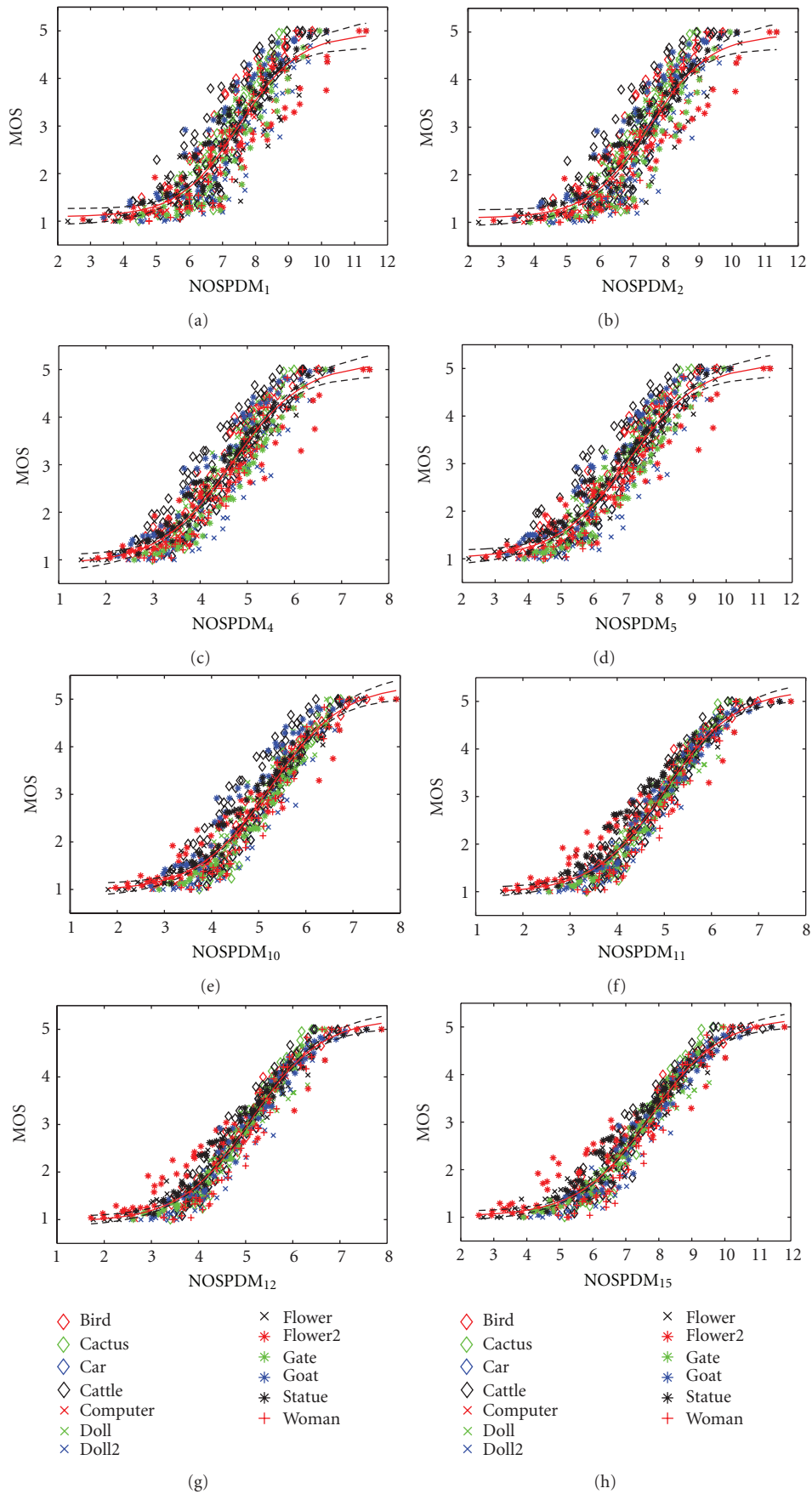


FIGURE 9: Scatter plots of MOS versus representative NOSPDM algorithms on the Toyama database [3]. The (red) lines are curves fitted with the logistic function and the (black) dash lines are 95% confidence intervals.

pair of images has the same JPEG compressive qualities (or lossless compression). Due to the consistent existence, this fact may be explained by the influence of inconsistent responses to different 3D image contents. According to a significant conclusion in [18] that the low subjective evaluation appeared for a high degree of parallax, and, moreover, based on a reasonable assumption that it is independent between parallax and the degree of image distortion (partly supported by the similar maps in Figure 7), it can be concluded that the uneven responses are caused by different degrees of parallaxes, which are introduced from different 3D image contents. For example, as shown in Figure 6, the “doll2” and “gate” images, which correspond to blue “×” and green “*” in Figure 4, have higher degrees of parallax, because of the content objects are closer to each other.

Thus, to balance this inconsistency, we first define the parallax as

$$d^\theta = \cos^{-1}\left(\frac{L \cdot R}{\|L\|_2 \cdot \|R\|_2}\right). \quad (7)$$

Furthermore, enlightened by VA based 2D image quality metrics [19, 20], a simple VA model defined in [21] is taken here to improve (7). Thus, the VA based parallax can be estimated by

$$d_V^\theta = \cos^{-1}\left(\frac{L(V) \cdot R(V)}{\|L(V)\|_2 \cdot \|R(V)\|_2}\right), \quad (8)$$

where V represents simple visual attention regions, as illustrated in Figure 5.

In addition, inspired by the behavior and the neuronal architectures of the early primate visual system, a classic bottom-up visual attention model (saliency model) [22] is to construct a single topographical saliency map first by combining multiscale image features, such as colors, intensity, orientations, and other visual information, and, then, a winner-take-all network that implements a neutrally distributed maximum detector is performed to detect the most salient locations step by step until the final saliency map is computed. The corresponding saliency maps of Figure 2 are displayed in turn in Figure 7. To further explore (7) by employing this saliency model, the saliency based parallax can be evaluated as

$$d_S^\theta = \cos^{-1}\left(\frac{L(S) \cdot R(S)}{\|L(S)\|_2 \cdot \|R(S)\|_2}\right), \quad (9)$$

where S represents saliency regions computed by the authors of [22].

Consequently, to compensate the low degree of parallax, the VA/saliency based parallax compensation can be computed by

$$F_{VP} = F_P + \omega_V \cdot d_V^\theta = d^\theta + \omega_V \cdot d_V^\theta \quad (10)$$

or

$$F_{SP} = F_P + \omega_S \cdot d_S^\theta = d^\theta + \omega_S \cdot d_S^\theta, \quad (11)$$

where ω_V and ω_S are model parameters. In the end, our proposed NOSPDM model is defined by

$$F_{NOSP} = F_O - F_N + F_{SP} = F_O^W - F_N - F_O^D + F_{SP}. \quad (12)$$

5. The Proposed Quality Metric

5.1. No-Reference IQA Method for JPEG Images. We aim to design a no-reference image quality metric for JPEG compressed stereoscopic image pairs. For JPEG compressed monoscopic images, at low bitrates, blurring and blocking artifacts occur due to the coarse quantization to 8×8 independent coding blocks. The blurring effect is mainly due to the loss of high frequency DCT coefficients while the blocking effect occurs because of the discontinuity at block boundaries.

Following [6], a 2D no-reference IQA metric is defined in four steps First, blockiness is estimated as the average differences across block boundaries as follows:

$$B_h = \frac{1}{M(\lfloor N/8 \rfloor - 1)} \sum_{i=1}^M \sum_{j=1}^{\lfloor N/8 \rfloor - 1} |d_h(i, 8j)|. \quad (13)$$

Second, the average absolute difference between in-block image samples is calculated as follows:

$$A_h = \frac{1}{7} \left[\frac{1}{M(N-1)} \sum_{i=1}^M \sum_{j=1}^{N-1} |d_h(i, j)| - B_h \right]. \quad (14)$$

Third, the horizontal zero crossing rate can be estimated as follows:

$$Z_h = \frac{1}{M(N-2)} \sum_{i=1}^M \sum_{j=1}^{N-2} z_h(m, n). \quad (15)$$

Finally, the image quality prediction is given by the following:

$$Q_{\text{JPEG}} = \alpha + \beta \left(\frac{B_h + B_v}{2}\right)^{\gamma_1} \left(\frac{A_h + A_v}{2}\right)^{\gamma_2} \left(\frac{Z_h + Z_v}{2}\right)^{\gamma_3}, \quad (16)$$

where B_v , A_v , and Z_v are the vertical features using similar methods like B_h , A_h , and Z_h , and α , β , γ_1 , γ_2 , and γ_3 are the model parameters.

5.2. NOSPDM Based Stereoscopic IQA Algorithm. Our proposed NOSPDM method primarily has five steps (only NOSPDM₂, NOSPDM₁₂, and NOSPDM₁₈ are illustrated in Figure 8 and other omitted NOSPDM algorithms have similar steps): first, predict the 2D image qualities of left and right images; second, apply different weighting coefficients to both image quality predictions; third, calculate the nonlinear quantity between them, based on nonlinear additive model or ocular dominance model; fourth, estimate VA/saliency based parallax compensation; fifth, evaluate the final stereoscopic image quality score by summing up different coefficients weighted 2D quality scores and parallax compensation and reducing the nonlinear quantity. Using Q_{JPEG} as an approximation to Q_{2D} , different combinations of NOSPDM _{x} ($x = 1 \dots 18$) tabulated in Table 3 are defined as follows:

$$\begin{aligned} \text{NOSPDM}_x &= F_L + F_O^W - F_N - F_O^D + F_P + F_{VP} + F_{SP} \\ &= Q_{\text{JPEG}}(L) + Q_{\text{JPEG}}(R) + F_O^W \\ &\quad - F_N - F_O^D + F_P + F_{VP} + F_{SP}. \end{aligned} \quad (17)$$

Finally, our proposed most effective NOSPDM₁₂ is given by

$$\begin{aligned}
\text{NOSPDM}_{12} &= F_O^W - F_N + F_{SP} \\
&= (2 - \mu_R) \cdot Q_{\text{JPEG}}(L) + \mu_R \cdot Q_{\text{JPEG}}(R) \\
&\quad - \lambda \cdot \max\{Q_{\text{JPEG}}(L), Q_{\text{JPEG}}(R)\} \\
&\quad + \cos^{-1}\left(\frac{L \cdot R}{\|L\|_2 \cdot \|R\|_2}\right) \\
&\quad + \omega_S \cdot \cos^{-1}\left(\frac{L(S) \cdot R(S)}{\|L(S)\|_2 \cdot \|R(S)\|_2}\right),
\end{aligned} \tag{18}$$

where λ , ω_V , and ω_S are model parameters. Here, we set λ equal to 0.67. This value should be explained by the ocular dominance theory, which needs more studies to further reveal the relationship between left- and right-eyes in the future. Besides, with respect to different reference images, we divide all the stereoscopic images in the Toyama database into two groups. Then, ω_V and ω_S can be determined through training on the first group (343 images) and testing on the second group (294 images).

5.3. Application of NOSPDM to Video Quality Metric. Video quality assessment (VQA) may be more relevant to real-world applications. The most salient video degradations include blurring, blockiness, and motion jerkiness artifacts. Blurring and blockiness can be measured in every single frame, but motion jerkiness has to be measured between successive frames. To measure the single frame quality, we notice that (16) can be rewritten as follows:

$$Q_{\text{JPEG}}(L) = Q_{\text{JPEG}}(L, \alpha, \beta, \gamma_1, \gamma_2, \gamma_3). \tag{19}$$

In this paper, NOSPDM applies the coefficients α , β , γ_1 , γ_2 , and γ_3 trained by [6] to obtain a stereoscopic IQA approach with high prediction performance. Then, according to [6], we know that blockiness is evaluated by the average differences across block boundaries, and average absolute difference between in-block image samples and zero crossing rate are employed to estimate the degree of blurring. So, it is believed that the single stereoscopic frame quality can be predicted through adjusting the above-mentioned coefficients. Thus, 3D video quality will be calculated by considering some temporal features of stereoscopic vision.

6. Experimental Results and Analyses

Mappings of these eighteen metrics values to subjective scores are obtained using nonlinear regression with a four-parameter logistic function as suggested by VQEG [25] as follows:

$$q(x) = \frac{\beta_1 - \beta_2}{1 + \exp(-x - \beta_3/\beta_4)} + \beta_2, \tag{20}$$

with x being the input score and $q(x)$ the mapped score and β_1 to β_4 are free parameters to be determined during the curve fitting process.

Five commonly used performance metrics as suggested by VQEG [25] are employed to further evaluate the competitive NOSPDM based stereoscopic IQA metrics on the Toyama database [3]. The first metric is the Pearson Linear Correlation Coefficient (PLCC) between MOS and the objective scores after nonlinear regression. It can be defined by

$$\text{PLCC} = \frac{\sum_i (q_i - \bar{q})(o_i - \bar{o})}{\sqrt{(\sum_i (q_i - \bar{q})^2)(\sum_i (o_i - \bar{o})^2)}}, \tag{21}$$

where o_i is the subjective score of the i th image. The second metric is the Spearman Rank-Order Correlation Coefficient (SRCC), computed as

$$\text{SRCC} = 1 - \frac{6\sum_{i=1}^N d_i^2}{N(N^2 - 1)}, \tag{22}$$

where d_i is the difference between the i th image's ranks in subjective and objective evaluations. It is a nonparametric rank-based correlation metric, independent of any monotonic nonlinear mapping between subjective and objective scores. The third metric, Kendall's Rank-Order Correlation Coefficient (KRCC), is another nonparametric rank correlation metric given by

$$\text{KRCC} = \frac{N_c - N_d}{(1/2)N(N - 1)}, \tag{23}$$

where N_c and N_d are the numbers of concordant and discordant pairs in the data set, respectively. Average Absolute Prediction Error (AAE) is the fourth metric, which is calculated using the converted objective scores after the nonlinear mapping of (20) the following:

$$\text{AAE} = \frac{1}{N} \sum |q_i - o_i|, \tag{24}$$

and the final metric Root Mean-Squared Error (RMSE) is defined by

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum (q_i - o_i)^2}. \tag{25}$$

All the values of different combinations of NOSPDM algorithms and [3] are presented in Table 4. And the representative scatter plots of NOSPDM methods are shown in Figure 9. Compared with image quality metrics of both the testing and training groups in the studies in [3], it is easy to find that our NOSPDM have achieved inspiring results and, moreover, as expected, the performance gain of NOSPDM₁₂ simultaneously based on nonlinear additive model, ocular dominance weighting model, and saliency based parallax compensation obtains the best performance.

Furthermore, through testing different combinations of NOSPDM methods illustrated in Table 3 and comparing the prediction accuracies in Table 4, some observations are given below.

Firstly, we can observe that stereoscopic image quality is considerably affected by nonlinear additive effects between

the qualities of left and right images, and through reducing the nonlinear quantity, the gain of PLCC and SRCC is about 0.04 (from NOSPDM_{1-2} to MOSPDM_{3-5}). Meanwhile, the saliency based parallax compensation is also very important with a gain of PLCC and SRCC around 0.035 (from MOSPDM_{3-5} to $\text{NOSPDM}_{9,12,15}$). Still, it is noticed that the ocular dominance weighting model can lead to a certain improvement, while its gain in PLCC and SRCC is just about 0.001 (from $\text{NOSPDM}_{1,3}$ to $\text{NOSPDM}_{2,4}$), and it is only valid for the nonlinear additive model. Besides, it is worth mentioning that the performance of ocular dominance model seems to be affected by the ocular dominance weighting model (from $\text{NOSPDM}_{5,13-15}$ to $\text{NOSPDM}_{6,16-18}$).

Secondly, from the prediction accuracy analysis in Table 4, it is easy to find that there is strong dependence between ocular dominance model and nonlinear additive model (e.g., $\text{NOSPDM}_{3,4}$, $\text{NOSPDM}_{10,13}$, etc.). And we can further observe from $\text{NOSPDM}_{4,10-12}$ to $\text{NOSPDM}_{5,13-15}$ that the integration of nonlinear additive model and ocular dominance weighting model is better than ocular dominance model only. Then, according to the illustrations in Figure 7 that left/right saliency maps of the four image pairs in Figure 2 are almost the same, and according to the fact that this phenomenon widely exists in stereoscopic images in the whole Toyama database, it can be concluded that the (saliency based) parallax is highly dependent on image contents, and the improved prediction accuracy is mainly based on reducing the influence of uneven responses to different image contents on perceptual stereoscopic image quality, which is supported by NOSPDM_4 to NOSPDM_{10-12} in Figure 9.

Eventually, some potential applications need to be emphasized here. In the utilization of NOSPDM , we note that the second most effective NOSPDM_{11} not only has more prediction performance than the studies in [3], but also has very little computational complexity, because of the simplicity of our used 2D no-reference IQA approach [6] and the proposed models, which only involves some basic computations, such as addition, subtraction, multiplication, and trigonometric functions. Meanwhile, some valuable applications are also illuminated in the area of JPEG image compression and stereoscopic video quality assessment. For the JPEG compression, under the condition of constant capacity of image storage, the approximately equal JPEG compressive qualities of left and right images tend to give higher subjective stereoscopic image quality. The related VQA approach can be further accomplished by adjusting some useful model parameters and incorporating some temporal features of stereoscopic vision, as stated in Section 5.3.

7. Conclusion

In this paper, we propose a novel no-reference stereoscopic image quality assessment algorithm based on 2D no-reference IQA method, nonlinear additive model, ocular dominance (weighting) model, and saliency based parallax compensation. By testing different combinations of NOSPDM algorithms, we have three key findings: firstly,

stereoscopic image quality is highly affected by nonlinear effects between left and right image qualities. Secondly, saliency based parallax compensation is quite important for improving prediction accuracy of image quality metrics. Finally, ocular dominance weighting model also contributes for the performance of the algorithm. By testing on the Toyama database, experimental results verify that our proposed NOSPDM methods have superior performance for stereoscopic image quality assessment.

It is natural to extend the stereoscopic IQA algorithms to stereoscopic video quality assessment. In the near future, we will construct a complete 3D video database. And moreover, due to the high prediction accuracy and low computational complexity of our NOSPDM approach, our work will be devoted to the research about stereoscopic video quality assessment, through taking into account some of the most salient video degradations, including blurring, blockiness, and motion jerkiness. Besides, we believe that the compressive consistency of stereoscopic JPEG images also warrants further study.

Acknowledgments

This work was supported in part by Postdoctoral Foundation of Shanghai 11R21414200, Postdoctoral Foundation of China 20100480603, 201104276, NSERC, NSFC (61025005, 60932006, 61001145), SRFDP (20090073110022), the 111 Project (B07022), and STCSM (12DZ2272600).

References

- [1] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images," *EURASIP Journal on Image and Video Processing*, vol. 2008, Article ID 659024, 2008.
- [2] A. Benoit, P. L. Callet, P. Campisi, and R. Cousseau, "Using disparity for quality assessment of stereoscopic images," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '08)*, pp. 389–392, October 2008.
- [3] Z. M. P. Sazzad, R. Akhter, J. Baltas, and Y. Horita, "Objective no-reference stereoscopic image quality prediction based on 2d image features and relative disparity," *Advances in Multimedia*, vol. 2012, Article ID 256130, 16 pages, 2012.
- [4] k. Gu, G. Zhai, X. Yang, and W. Zhang, "A new no-reference stereoscopic image quality assessment based on ocular dominance theory and degree of parallax," in *Proceedings of the IEEE International Conference on Pattern Recognition*, November 2012.
- [5] C. T. E. R. Hewage and M. G. Martini, "Reduced-reference quality evaluation for compressed depth maps associated with colour plus depth 3D video," in *Proceedings of the 17th IEEE International Conference on Image Processing (ICIP '10)*, pp. 4017–4020, September 2010.
- [6] Z. Wang, H. R. Sheikh, and A. C. Bovik, "No reference perceptual quality assessment of JPEG compressed images," in *Proceedings of the International Conference on Image Processing (ICIP '02)*, vol. 1, pp. I/477–I/480, September 2002.
- [7] H. R. Sheikh, A. C. Bovik, and L. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000," *IEEE Transactions on Image Processing*, vol. 14, no. 11, pp. 1918–1927, 2005.

- [8] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Processing Letters*, vol. 17, no. 5, pp. 513–516, 2010.
- [9] Z. Wang and A. C. Bovik, *Modern Image Quality Assessment. Synthesis Lectures on Image, Video & Multimedia Processing*, Morgan & Claypool, 2006.
- [10] J. C. A. Read, A. J. Parker, and B. G. Cumming, "A simple model accounts for the response of disparity-tuned V1 neurons to anticorrelated images," *Visual Neuroscience*, vol. 19, no. 6, pp. 735–753, 2002.
- [11] H. C. Nothdurft, "Saliency from feature contrast: additivity across dimensions," *Vision Research*, vol. 40, no. 10–12, pp. 1183–1201, 2000.
- [12] A. Macy, I. Ohzawa, and R. D. Freeman, "A quantitative study of the classification and stability of ocular dominance in the cat's visual cortex," *Experimental Brain Research*, vol. 48, no. 3, pp. 401–408, 1982.
- [13] M. Nomura, G. Matsumoto, and S. Fujiwara, "A binocular model for the simple cell," *Biological Cybernetics*, vol. 63, no. 3, pp. 237–242, 1990.
- [14] B. D. Chaurasia and B. B. L. Mathur, "Eyedness," *Acta Anatomica*, vol. 96, no. 2, pp. 301–305, 1976.
- [15] M. Reiss and G. Reiss, "Ocular dominance: some family data," *Laterality*, vol. 2, no. 1, pp. 7–15, 1997.
- [16] W. H. Ehrenstein, B. E. Arnold-Schulz-Gahmen, and W. Jaskhinski, "Eye preference within the context of binocular functions," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 243, no. 9, pp. 926–932, 2005.
- [17] I. Eser, D. S. Durrie, F. Schwendeman, and J. E. Stahl, "Association between ocular dominance and refraction," *Journal of Refractive Surgery*, vol. 24, no. 7, pp. 685–689, 2008.
- [18] S. Yano, S. Ide, T. Mitsunashi, and H. Thwaites, "A study of visual fatigue and visual comfort for 3D HDTV/HDTV images," *Displays*, vol. 23, no. 4, pp. 191–201, 2002.
- [19] H. Liu and I. Heynderickx, "Visual attention in objective image quality assessment: based on eye-tracking data," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 7, pp. 971–982, 2011.
- [20] A. Guo, D. Zhao, L. Shaohui, X. Fan, and W. Gao, "Visual attention based image quality assessment," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '11)*, pp. 3297–3300, September 2011.
- [21] Y. F. Ma, X. S. Hua, L. Lu, and H. J. Zhang, "A generic framework of user attention model and its application in video summarization," *IEEE Transactions on Multimedia*, vol. 7, no. 5, pp. 907–919, 2005.
- [22] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.
- [23] X. K. Yang, W. S. Lin, Z. Lu, E. P. Ong, and S. Yao, "Just-noticeable-distortion profile with nonlinear additivity model for perceptual masking in color images," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 609–612, April 2003.
- [24] J. Hoeve, "Amblyopia and squint," *Documenta Ophthalmologica*, vol. 7-8, no. 1, pp. 392–421, 1954.
- [25] VQEG, Final report from the video quality experts group on the validation of objective models of video quality assessment, 2000, <http://www.vqeg.org>.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

