Quality Index for Stereoscopic Images by Jointly Evaluating Cyclopean Amplitude and Cyclopean Phase

Yancong Lin, Jiachen Yang, Member, IEEE, Wen Lu, Member, IEEE, Qinggang Meng, Senior Member, IEEE, Zhihan Lv, and Houbing Song, Senior Member, IEEE

Abstract-With widespread applications of 3D technology, measuring quality of experience (QoE) for 3D multimedia content plays an increasingly important role. In this paper, we propose a full reference stereo image quality assessment (SIQA) framework which focuses on the innovation of binocular visual properties and applications of low-level features. On one hand, based on the fact that human visual system (HVS) understands an image mainly according to its low-level features, local phase and local amplitude extracted from phase congruency (PC) measurement are employed as primary features. Considering the less prominent performance of amplitude in IQA, visual saliency is applied into the modification on amplitude. On the other hand, by fully considering binocular rivalry phenomena, we create the cyclopean amplitude map and cyclopean phase map. With this method, both image features and binocular visual properties are mutually combined with each other. Meanwhile, a novel binocular modulation function in spatial domain is also adopted into the overall quality prediction of amplitude and phase. Extensive experiments demonstrate that the proposed framework achieves higher consistency with subjective tests than relevant SIQA metrics.

Index Terms—Stereo image quality assessment, Amplitude and phase, Phase congruency, Binocular visual properties, Human visual system

I. INTRODUCTION

T HE wave of 3D technology greatly enriches the way we perceive the world. With a large number of 3D films coming out and finally achieving commercial success, 3D technology has gained much more attention in creating a better psychological perception. The applications of 3D technology cover a wide range, including entertainments and scientific

Copyright (c) 2016 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org

This research is partially supported by the National Natural Science Foundation of China (No.61471260 and No.61271324) and Natural Science Foundation of Tianjin (16JCYBJC16000).

Y. Lin and J. Yang are with the School of Electronic Information Engineering, Tianjin university, Tianjin 30072, China (e-mail: yclin@tju.edu.cn; yangjiachen@tju.edu.cn).

W. Lu is with the School of Electronic Engineering, Xidian University, Xi'an, Shaanxi 710071, China (e-mail: luwen@mail.xidian.edu.cn).

Q. Meng is with the Department of Computer Science, School of Science at Loughborough University, UK (e-mail: q.meng@lboro.ac.uk).

Z. Lv is with the Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, 1068 Xueyuan Avenue, Shenzhen University Town, Shenzhen, 518055 China (e-mail: lvzhihan@gmail.com).

H. Song is with the Department of Electrical and Computer Engineering, West Virginia University, Montgomery, WV, 25136 USA (e-mail: Houbing.Song@mail.wvu.edu). research fields [1]–[5]. At the same time, we are also aware of the fact that all kinds of distortion types introduced during the process of transmission, processing and display will lead to the decrease of quality and great visual discomfort [6], [7]. Therefore designing an effective SIQA method is still a challenging area in image processing due to the complex and non-intuitive interactions of multiple 3D visual cues including depth perception, visual comfort and 2D-IQA related issues [8].

IQA mainly contains two aspects: subjective tests and objective tests. Compared with subjective tests which require numerous duplicate tests with a great number of participants [9], objective tests have great convenience and practicability, and achieved reliable prediction of perceived image quality. Therefore it attracts considerable attention of scholars and experts.

In practical applications, many 2D-IQA indexes largely depend on the availability of image features related to perceptual distortion, and then compute the overall quality from these features to depict the distribution of quality degradation, such as the Multi-Scale Structural Similarity Index Measurement (MS-SSIM) [10], and the Visual Information Fidelity (VIF) [11]. There are also a number of deep learning based blind image quality assessment (BIQA) frameworks (e.g., [12]–[14]) and natural scene statistics based IQA metrics which focus on the applications of quality-aware features (e.g., [15]–[18]).

However, it has been proved that 2D-IQA methods are not applicable to SIQA. You *et al.* investigated the appropriateness of ten well-known 2D-IQA metrics in SIQA [19]. The authors in [20] presented an evaluation of twelve 2D-IQA metrics to predict the quality of images synthesized by depth imagebased rendering (DIBR) [21]. The performance of these 2D-IQA metrics is moderate for measuring the quality of 3D images and they are not appropriate to assess the perceived depth, because of the weak correlation to stereo perception created by joint efforts of both views [8], [22].

Therefore, how to get a full understanding of the perceptual distortion on stereopairs under the assumption of various binocular visual properties is still the focus of SIQA [23], [24]. For example, the two views of the stereoscopic images may contain different distortion types due to the asymmetric processing in the 3DTV system such as capture, coding and transmission. The asymmetric distortion in stereoscopic images exists with different intensity, locations and texture structures. Then the binocular fusion or rivalry may occur in

different locations [25]. With this purpose, both the availability of distortion-relevant image features and binocular visual models need further analysis.

On one hand, the new challenges of SIQA come mainly from the interactions between the two eyes. Aiming at a better simulation of cyclopean images, Levelt has proved that binocular rivalry/suppression is mainly governed by low-level sensory factors, and stimuli that are higher in contrast tend to dominate the rivalry [26]. Maalouf et al. integrated the left and right views into a cyclopean image, by simulating the brain perception, and used the contrast sensitivity coefficients as the basis of evaluation [27]. In subsequent work, Chen et al. proposed a full reference metric which focused on the highly relevant phenomena of binocular rivalry [28]. Lin et al. applied binocular frequency integration into measuring the quality of stereoscopic images, with the help of traditional 2D metrics [24]. In [29], a new full-reference SIOA framework was proposed by learning binocular receptive field properties in line with visual perception. These metrics do quite well in predicting perceptual quality of stereo images. Inspired by these positive results, a synthesized framework which takes binocular rivalry into consideration is proposed.

On the other hand, the HVS understands an image mainly based on its low-level features, such as edges and zero crossings [30], [31]. Accordingly, perceptible image degradation will lead to perceptible changes in image low-level features. As illustrated in [32], image features such as step edges, lines and Mach bands all give rise to the points where the Fourier components of an image are maximal in phase. Inspired by this conclusion, Phase Congruency (PC) measurement is employed into practical applications.

The PC model is a frequency-based model of visual processing. It supposes that, instead of processing visual data spatially, the visual system is capable of performing calculations using the phase and amplitude of the individual frequency components in a signal [32]. In [33], Liu and Laganière put a local cross-correlation of PC map into the calculation between a reference image and a distorted one, and the averaged crosscorrelation values provided a quantitative assessment of the overall image similarity. In [34], PC was extended to phase coherence which can be used to characterize the image blur. Based on [34], Hassen et al. proposed a sharpness measure where sharpness was identified as strong local phase coherence [35]. In [22], Shao et al. applied the PC calculation together with the log-Gabor filter into the description of local amplitude and phase features. However, amplitude can not reflect as much detailed information as phase through the analysis on image construction, and therefore receives a lower weighting coefficient. The authors in [36] carried out experiments to investigate the visual impact of distortion on phase and amplitude of an image. It was found that the distortion in phase had a bigger impact on the visual appearance of image and its quality [37].

It seems to be a general consensus that phase plays a more prominent role in characterizing the signal properties. However, to highlight the importance of amplitude in determining visual quality, we integrate amplitude with visual saliency information.

When catching sight of an image, our eyes attempt to recognize certain regions perceptually important in order to speed up the recognition process and to extract the significant objects from complex backgrounds [38]. Saliency detection comes into being exactly in accordance with this thought [39]-[43]. There are several saliency detection models succeeding in highlighting the locations that attract human attention, while weakening the parts that HVS is insensitive to, such as the widely accepted bottom-up method named Graph-Based Visual Saliency (GBVS) [39], the saliency detection model using Maximum Symmetric Surround (MSS) [40], the model using Low-level Features based on Wavelet Transform (LFWT) [41] and the model based on the Image Signature (IS) [42]. Zhang et al. used visual saliency as primary features when computing the local quality map of distorted images [3]. Furthermore, visual saliency can be employed as a weighting function to reflect the importance of a local region in the pooling procedure [8]. At the same time, the emerging applications of stereoscopic display require new saliency detection models for salient region extraction. Fang et al. took the depth feature into account in saliency detection for stereoscopic images and proposed a saliency detection model based on the feature contrast of color, luminance, texture, and depth [44]. According to this characteristic, visual saliency is adopted into the modification on amplitude to bridge the gap between amplitude and visual quality.

Inspired by previous work, a full reference SIQA framework based on the cyclopean amplitude map and the cyclopean phase map is proposed. The contributions are as follows:

1) Considering the less prominent role of amplitude in IQA, visual importance map which contains efficient saliency and depth information, is creatively adopted into the modification on amplitude maps.

2) Local amplitude (phase) maps extracted from PC measurement of both views are synthesized as one cyclopean amplitude (phase) map which represents the perceptual degradation of stereopairs after binocular rivalry process.

3) By comparing the matching error in binocular disparity, stereoscopic images are classified into non-corresponding and corresponding regions, and a modulation function in spatial domain is applied into the quality prediction.

4) Through comprehensive experiments and comparison, we analyze the contribution of amplitude and phase in determining visual quality. Results show that our metric correlates well with subjective tests on both symmetrically and asymmetrically distorted stereoscopic images.

The rest of this paper is organized as follows. Section II illustrates the related work of creating cyclopean phase and amplitude maps, including image features and binocular visual properties. Section III presents the detailed computation procedures of proposed SIQA framework. The experimental results are analyzed in Section IV. Finally in Section V, we conclude the paper with the imagination of future work.

II. RELATED WORK

A. Definition of the Cyclopean Phase Map and Cyclopean Amplitude Map

The purpose of cyclopean image is to give an intuitive description of the perceived image acquired from both left and right views. To achieve this goal, we have to extract useful features that can effectively describe visual degradation. As mentioned in Introduction, the phase and amplitude maps can also be considered as perceptual representations from different aspects, and therefore the integration between the two maps and the cyclopean image is practicable. With this method, we get a pair of cyclopean maps named cyclopean phase map and cyclopean amplitude map. On one hand, the cyclopean phase map plays the dominant role in expressing details of the perceived image (e.g., structure information), and thus it is reasonable to be taken into consideration. On the other hand, the cyclopean amplitude map is deemed as another evaluation factor, owing to the approximate reflection of visual saliency after the integration with visual importance map. The detailed information about modification on amplitude is given in the following part. Figure. 1 shows examples of cyclopean phase and cyclopean amplitude maps.



Fig. 1: An examples of cyclopean phase and cyclopean amplitude maps. (a) Original left image. (b) Original right image. (c) The cyclopean phase map. (d) The cyclopean amplitude map. Note that all the original images come from LIVE 3D Image Quality Database [45]

B. Modification on Amplitude

As shown in Fig. 2, there are several constructed images using the phase of reference images, and amplitude of JPEG compressed, Gaussian blurred and fast faded images. From the comparison among the original reference image and constructed images, it can be observed that the constructed images are degraded due to compression and blurring on the amplitude, despite most details are preserved. Therefore amplitude can also be used to reflect image degradation. However, the connection between amplitude and visual quality is still less prominent [22], [36], [37]. To solve this problem, we integrated the amplitude with a saliency-based visual importance map estimation (shown in Fig. 3) which did an excellent job in reflecting certain regions with different visual importance [6]. With this method, certain regions and objects that are perceptually important in visual perception will be highlighted in modified amplitude maps (shown in Fig. 4), such as the parking space in image "Bicycle" and the ground in image "Park". In contrast, the original amplitude shows less useful information.



Fig. 2: Examples of image reconstruction using the phase of original left reference image "Park" and amplitude of corresponding distorted images. (a) Original left reference image. (b) Image constructed using the amplitude of JPEG compressed image. (c) Image constructed using the amplitude of blurred image. (d) Image constructed using the amplitude of fast faded image.

The detailed estimation procedure of visual importance map is shown in Fig. 5. As an example, the GBVS model which highlights low-level features in monocular view is employed to compute the saliency map (S_I). Regarded as the crucial difference to 2D IQA, disparity is used to compute the depth saliency map (S_D) after normalization [46], [47]. Finally, the image saliency map and depth saliency map are integrated as the visual importance map (S) by linear combination given by $S = \lambda S_I + (1-\lambda)S_D$, where λ denotes the weighting factor of the saliency map (in our experiment, $\lambda = 0.5$). The examples of synthesized visual importance maps are shown in Fig. 3.

C. Binocular Spatial Modulation Function

The binocular suppression and rivalry properties occur when the two views of stereoscopic images fall on the corresponding retina locations. Therefore, each pixel in either view should be divided into Corresponding Regions (CR) and Non-Corresponding Regions (NCR), according to whether the pixels have correspondences in the other view [25]. A novel modulation function in spatial domain is adopted into the pooling by considering the binocular suppression and rivalry properties. The Binocular Spatial Modulation Function (M) is implemented in corresponding regions and non-corresponding regions according to the disparity map, and it can be defined as

$$M(x,y) = \begin{cases} \phi((S(x,y), S(x,y)/JND(x,y)), & (x,y) \in NCR\\ \phi((S(x,y), S(x,y)/BJND(x,y)), & (x,y) \in CR \end{cases}$$
(1)

IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING



Fig. 5: Estimation procedure of Visual Importance Map introduced in [6].



Fig. 3: Examples of visual importance map estimation. (a) Original left image "Bicycle". (b) Original left image "Park". (c) Saliency map of (a). (d) Saliency map of (b). (e) Depth saliency map of (a). (f) Depth saliency map of (b). (g) Visual importance map of (a). (h) Visual importance map of (b). Note that, in (g) and (h), visual importance values range from 0 (least importance, blue) to 1 (most importance, red).



Fig. 4: Modification on amplitude maps. (a) Original left image "Bicycle". (b) Original left image "Park". (c) Amplitude map of (a). (d) Amplitude map of (b). (e) Modified amplitude map of (a). (f) Modified amplitude map of (b).

where JND and BJND denote the existing just-noticeable difference model [48] and the binocular just-noticeable difference model [49] in NCR and CR, S denotes the visual importance map, and ϕ is a nonlinear additive function [25]. Here, the function ϕ is computed by

$$\phi(x,y) = (x+y)/2 - \kappa \cdot \min(x,y) \tag{2}$$

where κ is a parameter adjusting the penalization strength. In the proposed framework, it is empirically set to 0.45 for all types of distortion [50].

(B)JND means the tolerance for distortion, and thus it can determine the minimum distortion that evoke binocular

visible differences [49]. Intuitively, if a region has a significant (B)JND value, it implies that the region can tolerate large distortion, and importance of the region in the binocular perception is low [50].

In summary, image features and binocular visual properties analyzed in this section have major impacts on SIQA: 1) the amplitude map allows better representation of image features, and provides the basic identification of visual attention. 2) the cyclopean amplitude and phase maps give an intuitive description of visual degradation. 3) the binocular spatial modulation function can also be used into SIQA for more effective evaluation.

III. PROPOSED SIQA FRAMEWORK

The proposed SIQA framework mainly contains three sequential parts: the extraction of local amplitude and phase maps, the computation of cyclopean amplitude and phase maps, and the pooling process, as shown in Fig. 6. Firstly, the local amplitude and phase maps are extracted from the reference stereopairs, as well as the distorted stereopairs; Secondly, local amplitude maps between two views are synthesized as one cyclopean amplitude map, and the same is true of the local phase maps; At last, the similarities between two cyclopean phase maps as well as two cyclopean amplitude maps are applied into the pooling process respectively, and then a linear regression is employed to combine the two similarity indexes into the final quality. In order to facilitate understanding, the summary of some important notations and variables is shown in Table I.

A. Local Amplitude and Local Phase Maps Extracted from PC

As analyzed above, the modified amplitude and phase maps are important features in IQA. However, global phase from the direct Fourier transform is not an effective means in feature description, because different regions are evaluated independently. Obviously, the constant frequency response is not a good choice for feature description [22]. According to the phase congruency theory, the local phase and amplitude features can be extracted by maximizing the Fourier components [51].

Morrone and Owens [52] proposed an image feature perception model, which postulated that features were perceived at the points where Fourier components were maximal in phase [33]. Kovesi proposed a scheme to calculate the phase congruency with logarithmic Gabor wavelets, which allows a wide range of features to be detected within the framework of a single model [32]. Inspired by these conclusions, PC measurement is employed into practical applications. It has been proved that simple cells in primary visual cortex can be well modeled using log-Gabor filter in previous researches [53]. A set of responses on different scales and along different orientations, denoted as $[\eta_{s,o}, \xi_{s,o}]$, can be obtained by applying the log-Gabor filter $G_{s,o}$ (denoted by spatial scale index *s* and orientation index *o*) in Fourier domain:

$$G_{s,o}(\omega,\theta) = exp[-\frac{(log(\frac{\omega}{\omega_s}))^2}{2\sigma_s^2}] \cdot \exp[-\frac{(\theta-\theta_o)^2}{2\sigma_o^2}]$$
(3)

where parameters ω and θ are the normalized radial frequency and the orientation angle of the filter, and ω_s and θ_o are the corresponding center frequency and orientation of the filter, respectively. The parameters σ_s and σ_o determine the strength of the filter. Then the local amplitude at location x on scale s and along orientation o is given by

$$A_{s,o}(x) = \sqrt{\eta_{s,o}(x)^2 + \xi_{s,o}(x)^2}$$
(4)

and the local energy along orientation o is computed by

$$E_o(x) = \sqrt{F_o(x)^2 + H_o(x)^2}$$
 (5)

where $F_o(x) = \sum_s \eta_{s,o}(x)$ and $H_o(x) = \sum_s \xi_{s,o}(x)$. The PC along orientation *o* is computed by

$$PC_o(x) = \frac{E_o(x)}{\varepsilon + \sum_s A_{s,o}(x)} \tag{6}$$

where ε is a small positive constant. The design of the filter is based on the work in [54].

Different with [55], the PC is directly used into feature description. The local phase is defined as the angle of $F_o(x)$ and $H_o(x)$ along the orientation with the maximum PC value:

$$LP(x) = \arctan(H_{o_m}(x), F_{o_m}(x)) \tag{7}$$

where o_m denotes the orientation which corresponds to the maximum PC value. Then the local amplitude is defined as the sum of local amplitudes of all the scales along the orientation o_m :

$$LA(x) = \sum_{s} A_{s,o_m}(x) \tag{8}$$

According to these definitions, local amplitude and phase maps extracted from original images and corresponding distorted versions are displayed in Fig. 7 and Fig. 8. In order to facilitate visual examination, those regions with significant changes are marked in red rectangles. As shown in Fig. 7, distortion leads to visual degradation on key objects (e.g., the bicycle). There are also similar effects on local phase maps which describe totally different visual contents from local amplitude maps, for example, the structure information is preserved.

In our experiment, $\omega_s = 1/6$, $\theta_o = 0$, $\sigma_s = 0.3$, $\sigma_o = 0.4$, and the numbers of scale and orientation of the filter are set to 4.

B. Cyclopean Amplitude and Phase Maps

Figure. 9 shows how to compute the cyclopean phase and amplitude maps. At the beginning, an image from either view is decomposed into local phase and amplitude maps, and visual importance map is also acquired. Secondly, synthesized by corresponding visual importance map, the modified amplitude maps are applied into the following procedures. Eventually, the modified amplitude maps from two views are combined into a cyclopean amplitude map, and the same is true of the two phase maps. Both reference and distorted stereopairs are applied into the computation.



Fig. 6: Procedure of the proposed method.



S_I	the saliency map	LP, LA	the local phase and local amplitude maps
S_D	the depth saliency map	A_L, A_R	the modified amplitude maps of left and right images
S	the visual importance map	G	the complex 2D Gabor filter
λ	the weighting factor of the saliency map	GE_L, GE_R	the summation of convolution responses
M	the binocular spatial modulation function		of the left and right images to G
(B)JND	the (binocular) just noticeable difference	CP, CA	the cyclopean phase and cyclopean amplitude
(N)CR	the (Non-) Corresponding Region	CM	the cyclopean expression of M
ϕ	the function to compute M	W_L, W_R	the left and right weighting factors
$G_{s,o}$	the log-Gabor filter on scale s and along orientation o		computed from GE_L , GE_R
ω	the normalized radial frequency	S_A, S_P	feature similarity metrics for CP, CA
θ	the orientation angle of the filter	Q_A, Q_P	the quality metrics of amplitude and phase
$A_{s,o}$	the local amplitude on scale s and along orientation o	γ	the weighting factor for cyclopean phase
E_o	the local energy along orientation o	Q	the final quality
PC_o	PC along orientation o		



Fig. 7: Examples of local amplitude maps extracted from original left image "Bicycle" and corresponding distorted versions (white noise and fast fading, respectively).



Fig. 8: Examples of local phase maps extracted from the original left image "Bicycle" and corresponding distorted versions (white noise and fast fading, respectively).



Fig. 9: Procedure of the cyclopean phase and amplitude computation.

The process of synthesized cyclopean phase and amplitude maps is illustrated in the following part. According to [26], [28], the energy of Gabor filter bank responses on the left and right images can be used to model stimulus strength and simulate rivalrous selection of cyclopean image quality. In general, the complex 2D Gabor filter can be defined as

$$G(x, y, \sigma_x, \sigma_y, \zeta_x, \zeta_y, \theta) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-(1/2)[(R_1/\sigma_x)^2 + (R_2/\sigma_y)^2]e^{i(x\zeta_x + y\zeta_y)}}$$
(9)

where $R_1 = x \cos \theta + y \sin \theta$ and $R_2 = y \cos \theta - \sin \theta$, σ_x and σ_y are the standard deviations of an elliptical Gaussian envelope along x and y axes, ζ_x and ζ_y are spatial frequencies, and θ orients the filter. The design and parameter settings of the Gabor filter are derived from the work conducted by Su *et al.* [56]. The local energy is estimated by summing Gabor filter magnitude responses over four orientations (horizontal, both diagonals, and vertical) at a spatial frequency of 3.67cycles/degree [28].

Given the Gabor filter response of each view, a linear model is used to synthesize the cyclopean phase and amplitude maps:

$$CP(x,y) = W_L(x,y) \times LP_L(x,y) + W_R((x+d),y) \times LP_R((x+d),y)$$
(10)

$$CA(x,y) = W_L(x,y) \times A_L(x,y) + W_R((x+d),y) \times A_R((x+d),y)$$
(11)

where CP and CA represent the cyclopean phase map and the cyclopean amplitude map, A_L and A_R are modified local amplitude maps of the left and right image, and the same is true of local phase maps LP_L and LP_R . The weights $W_L(x, y)$ and $W_R(x + d, y)$ are computed from the normalized Gabor filter magnitude responses computed by

$$W_L(x,y) = \frac{GE_L(x,y)}{GE_L(x,y) + GE_R((x+d),y)}$$
(12)

$$W_R((x+d), y) = \frac{GE_R((x+d), y)}{GE_L(x, y) + GE_R((x+d), y)}$$
(13)

where GE_L and GE_R are the summation of convolution responses of the left and right images to the complex 2D Gabor filter.

Different from the model proposed by Levelt [26], the cyclopean image is disparity-compensated in computation equation, denoted as (x + d). Accordingly, the cyclopean map is able to reconstruct the perceptual image that matches HVS to a large extent by taking binocular rivalry into consideration.

C. Quality Assessment Metric

Figure. 6 shows how the proposed framework works after cyclopean phase and amplitude computation. Similar to the definition of SSIM [57], the two feature similarity metrics for CP and CA between reference stereopairs and distorted pairs are adopted, denotes as S_A and S_P , which can be computed as follows:

$$S_A = \frac{2CA_{\text{ref}} \cdot CA_{dis} + T_1}{CA_{\text{ref}}^2 + CA_{dis}^2 + T_2} \tag{14}$$

$$S_{P} = \frac{2CP_{\text{ref}} \cdot CP_{dis} + T_{1}}{CP_{\text{ref}}^{2} + CP_{dis}^{2} + T_{2}}$$
(15)

where T_1 and T_2 are positive constants to avoid the denominator being zero. The subscripts ref and dis represent reference and distorted stereopairs respectively. After computing the two similarity metrics of cyclopean maps at each location, the binocular spatial modulation function is applied into the pooling procedure. More specifically, the general mathematics forms of quality assessment metric for amplitude and phase are given by

$$Q_A = \frac{\sum_{x,y} CM(x,y) \cdot S_A(x,y)}{\sum_{x,y} CM(x,y)}$$
(16)

$$Q_P = \frac{\sum_{x,y} CM(x,y) \cdot S_p(x,y)}{\sum_{x,y} CM(x,y)}$$
(17)

where Q_A and Q_P are the amplitude quality metric and phase quality metric respectively, and CM(x, y) denotes the cyclopean binocular spatial modulation function based on distorted stereopairs, which can also be defined as

$$CM(x,y) = W_L(x,y) \times M_L(x,y) + W_R(x+d,y) \times M_R(x,y)$$
(18)

Finally, in order to get the final quality Q, a linear combination between Q_P and Q_A is employed, given by

$$Q = \gamma Q_P + (1 - \gamma) Q_A \tag{19}$$

where γ is the weighting factor for adjusting the relative importance of the two components. Analysis on how the modified amplitude and phase maps influence the quality assessment under the circumstances of different weights will be presented in Section IV.

IV. EXPERIMENTAL RESULTS AND ANALYSES

A. Stereo Database

To verify the performance of the proposed method, the LIVE 3D Image Quality Databases (Phase I and Phase II) of the University of Texas at Austin are used [45]. Database Phase I database contains 365 stereopairs with symmetric distortion and Database Phase II contains 360 stereopairs with both asymmetric and symmetric distortion, including JPEG compression, JP2K compression, white noise (WN), Gaussian blur(Blur) and fast fading(FF).

To further verify the proposed method, the IRCCyN 3D Image Database is also adopted [58]. Six different stereoscopic images are considered in this database which is composed of the six reference images (undistorted) and fifteen distorted versions of each source generated from three different processings (JPEG compression, JP2K compression and blurring) symmetrically to the stereopairs.

B. Performance Measure

For performance evaluation, four commonly used indicators are adopted: Pearson Linear Correlation Coefficient (PLC-C), Spearman Rank Order Correlation Coefficient (SROCC), Kendall Rank-order Correlation Coefficient (KROCC), and Root Mean Squared Error (RMSE) between subjective scores and objective scores after nonlinear regression. For nonlinear regression, we use a 4-parameter logistic mapping function [23], [59]:

$$DMOS_P = \frac{\beta_1 - \beta_2}{e^{(Q - \beta_3)/|\beta_4|} + 1} + \beta_2$$
(20)

where β_1,β_2,β_3 and β_4 are the parameters to be fitted. A better match is expected to have higher PLCC, SROCC, and KROCC values, while lower RMSE values.

In order to demonstrate its efficiency, the proposed method is compared with several existing state-of-art IQA metrics, including three 2D metrics (MS-SSIM [10], VIF [11] and ADM [60]), and four 3D metrics (Benoit's scheme [61], Lin's scheme [24], Chen's scheme [28], Shao's scheme-A [22] and Shao's scheme-B [29]). For Benoit's scheme, we adopt the d_2 metric, in which SSIM is used and disparity distortion is estimated using global correlation coefficient between the original and disparity maps. For Lin's scheme, the FI-PSNR metric is adopted into the comparison. For Chen's scheme, the adopted 2D metric is MS-SSIM which performs the best. For Shao's schemes, experimental results in corresponding references are directly adopted.

C. the Joint Contribution of Cyclopean Amplitude Map and Cyclopean Phase Map

To better illustrate the relationship between the cyclopean amplitude map and cyclopean phase map under the circumstances of different weights, the prediction curves on LIVE 3D Image Database are shown as an example in Fig. 10. According to the fitting curves of Database Phase I, with the increasing weight of cyclopean phase map (γ), the prediction indicators seem to drop, except for RMSE. At the beginning of the fitted curve ($\gamma = 0$), the cyclopean amplitude map is given a full weight and the four prediction indicators perform quite well. In the end, when the cyclopean phase map is given a full weight ($\gamma = 1$), the performance of four prediction indicators are not so excellent as others. The difference reconfirms that the cyclopean amplitude takes effect in SIQA, since the visual saliency information is taken into account in the process of amplitude modification introduced in Section II. Although there are minor differences in numeric values, the other saliency detection models also show similar performances. According to the fitting curves, when γ equals to 0.60, the four indicators achieve more accurate results.

However, the corresponding prediction curves on LIVE 3D Image Database-Phase II show a relatively different trend from the other one. There are significant peaks in PLCC, SROCC and KROCC, and a valley in RMSE at the position close to γ =0.75, which indicates that the cyclopean phase map still plays a leading role.

Comparing the two groups of fitting curves in Fig. 10, there are some difference between them. The prediction results on symmetric distortion are more stable with varying γ , and in contrast, remarkable changes can be observed in the case of Phase II. This may be explained by the following example. As shown in Fig. 11, the modified amplitude maps of a pair of asymmetrically distorted stereo images (fast fading and JP2K compression, respectively) and the corresponding cyclopean amplitude map are compared with the reference images. When it comes to asymmetric distortion, the distortion



Fig. 10: The joint contribution of cyclopean amplitude and cyclopean phase on LIVE 3D Image Databases under the circumstances of different weights. Note: Blue fitting curves represent the results on Database Phase I, and red curves represent the results on Database Phase II.

introduced by fast fading and JP2K compression are different from each other in both locations and distortion levels (marked in the red rectangle), and details can also be found in the cyclopean amplitude map. Therefore, the quality assessment of asymmetrically distorted images is a relatively difficult task considering the joint effects of various factors. In this case, only depending on the cyclopean amplitude maps to reflect visual degradation would not be able to achieve a better consistency with subjective assessment. As shown in Fig. 10, the prediction curves on Database Phase II show a relatively poorer performance at the beginning (e.g., $\gamma=0$). In addition, the phase maps (shown in Fig. 8,) which describe totally different visual contents from the local amplitude maps, could be regarded as another factor to enhance the quality assessment. According to this, the prediction curves show remarkable changes and peak at nearly γ =0.75. However, in terms of symmetric distortion, images are degraded in a similar degree in both views, and predicting the quality would be comparatively easy. Thus the contribution of phase is less significant and minor changes can be found on Database Phase I.

In order to get a more general quality index for 3D images, setting a reasonable parameter for the two components is of vital importance. Considering the performances on two databases, the parameter is adjusted as $\gamma = 0.75$, since the corresponding performances on Database Phase I are still competitive. In the following comparisons, the parameter in the proposed method is $\gamma = 0.75$.

To demonstrate the advantage of jointly considering cyclopean amplitude and cyclopean phase, Table II presents the detailed results of the following scenarios:

1) Only cyclopean amplitude is used ($\gamma = 0$);

2) The best results on LIVE 3D Image Database Phase I ($\gamma = 0.6$);

3) The best results on LIVE 3D Image Database Phase II $(\gamma = 0.75)$;

4) Only cyclopean phase is used ($\gamma = 1$);

The scatter plots on LIVE 3D Image Database are shown in Fig. 12. The vertical axis denotes the human subjective scores of perceptual quality and the horizontal axis denotes the predicted scores. A better convergence of the points in the scatter plot to the fitting curve means a better consistency with the DMOS. The scatter plots demonstrate that the scatter points of the proposed method are concentrated to the fitting curve, which indicates a more prominent prediction.

D. Overall Performance on LIVE 3D Image Database

To better illustrate the experimental results, the values of PLCC, SROCC, KROCC, and RMSE of each metric are shown in Table III, where the indicator that gives the best performance is highlighted in bold. At the same time, results of the four saliency detection models adopted in the proposed framework (GBVS, MSS, LFWT, and IS) are also shown in Table III. As shown in the table, the four saliency detection models show similar performances with each other, and the overall performances of the proposed framework significantly better compared with other IQA metrics on both databases. Although in some cases, their performances slightly lag behind the best, the overall prediction results are still competitive among all the IQA metrics. The outstanding performance partially demonstrates the reasonability of cyclopean amplitude and cyclopean phase maps.

Although Lin's scheme considered binocular integration behaviors, the lack of disparity information may lead to the unexpected performance. Known as the combination of 2D image quality metrics with disparity maps, Benoit's scheme still lags behind the other metrics. In view of this, it can be inferred that the quality of the estimated disparity maps highly depends on the stereo matching algorithms and the 2D-IQA metric for disparity maps does not coincide with human perception of disparity [29]. As for Shao's schemes, they achieve reasonably accurate prediction results on Database Phase I, while the performances on Database Phase II are not so excellent as Chen's scheme and the proposed schemes.

In addition to symmetric distortion, the following part concentrate more on asymmetric distorted stereopairs. According to the experimental results on Database Phase II which contains both symmetric and asymmetric distortion, the proposed framework and Chen's scheme overtake the other metrics by a large extent. The significant difference further confirms the previous conclusion that cyclopean images can effectively predict the quality of stereoscopically viewed images. However, it is worth noting that predicting the quality of asymmetric distorted stereopairs still needs further improvements, compared with symmetric distortion.

In spite of the good performance of overall quality prediction, we also need to verify the performance of each



Fig. 11: Modified amplitude maps from the left and right views, and corresponding cyclopean amplitude maps of the stereopair "Bicycle". Figures in the first row come from the original stereopair; figures in the second row come from the asymmetrically distorted stereopair (JP2K compression for the left view and fast fading for the right view).

TABLE II: The joint contribution of cyclopean amplitude map and cyclopean phase map

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	LIVE 3D Image Database Phase I				LIVE 3D Image Database Phase II			
Ĭ	PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE
$\gamma = 0$ (cyclopean amplitude only)	0.9294	0.9320	0.7654	6.047	0.8553	0.8186	0.6526	5.848
$\gamma = 0.6$	0.9388	0.9347	0.7687	5.647	0.904	0.8887	0.7115	4.825
$\gamma = 0.75$	0.9366	0.9314	0.7631	5.744	0.9113	0.8935	0.7206	4.648
$\gamma = 1$ (cyclopean phase only)	0.9059	0.9001	0.7157	6.946	0.8831	0.8683	0.6863	5.296

## TABLE III: Overall performances on LIVE 3D Image Database

Criteria		LIVE	E 3D Image	Database Pl	hase I	LIVE 3D Image Database Phase II					
	Cincila	PLCC	SROCC	KROCC	RMSE	PLCC	SROCC	KROCC	RMSE		
	MS-SSIM [10]	0.9252	0.9233	0.7474	6.2245	0.7758	0.7707	0.6023	7.1221		
	VIF [11]	0.9117	0.9087	0.7220	6.7368	0.7865	0.7167	0.5427	6.9708		
	ADM [60]	0.9119	0.9083	0.7197	6.7314	0.7513	0.7242	0.5481	7.4496		
	Benoit [61]	0.8015	0.7936	0.5928	9.8058	0.7423	0.7258	0.5357	7.5632		
	Lin [24]	0.8645	0.8559	0.6559	8.2424	0.6584	0.6375	0.4701	8.4956		
	Chen [28]	0.9161	0.9153	0.7360	6.5740	0.9067	0.9068	0.7314	4.7587		
	Shao-A [22]	0.9245	0.9217	-	6.2522	0.7585	0.7451	-	7.3554		
	Shao-B [29]	0.9350	0.9251	-	5.8155	0.8628	0.8494	-	5.7058		
	Proposed-GBVS	0.9347	0.9295	0.7589	5.8299	0.8967	0.8805	0.7044	4.9954		
	Proposed-MSS	0.9324	0.9276	0.7563	5.9253	0.9031	0.8867	0.7125	4.8470		
	Proposed-LFWT	0.9366	0.9314	0.7631	5.7440	0.8984	0.8824	0.7057	4.9565		
	Proposed-IS	0.9242	0.9203	0.7415	6.2629	0.9113	0.8935	0.7206	4.6477		

TABLE IV: Detailed performances on LIVE 3D Image Database Phase I

IADLE IV. Detailed performances on LIVE 3D image Database r mase I										
Distortion	Criteria	MS-SSIM [10]	VIF [11]	ADM [60]	Benoit [61]	Lin [24]	Chen [28]	Shao-A [22]	Shao-B [29]	Proposed
JP2K	PLCC	0.9188	0.9373	0.9333	0.7627	0.8381	0.9166	0.9238	0.9213	0.9520
	SROCC	0.8917	0.9015	0.9008	0.7527	0.8388	0.8954	0.8752	0.8945	0.9127
	KROCC	0.7006	0.7139	0.7158	0.5519	0.6386	0.7133	-	-	0.7468
	RMSE	6.8141	6.0168	6.2009	8.3767	7.0658	5.1771	-	-	3.9631
IEDC	PLCC	0.6859	0.6779	0.6719	0.2164	0.2866	0.6356	0.6563	0.5200	0.7546
	SROCC	0.6123	0.5807	0.5926	0.0754	0.0931	0.5632	0.6148	0.4951	0.7164
JEIO	KROCC	0.4179	0.3908	0.4009	0.0507	0.2140	0.3749	-	-	0.5161
	RMSE	6.3445	6.4097	6.4578	6.3868	6.2650	5.0486	-	-	4.2912
	PLCC	0.9320	0.9203	0.9211	0.9256	0.9280	0.9353	0.9410	0.9448	0.9266
WN	SROCC	0.9320	0.9221	0.9235	0.9241	0.9284	0.9376	0.9431	0.9405	0.9289
VV IN	KROCC	0.7753	0.7519	0.7608	0.7595	0.7614	0.7880	-	-	0.7671
	RMSE	8.0413	8.6781	8.6342	6.2971	6.1964	5.8874	-	-	6.2569
	PLCC	0.9435	0.9598	0.9607	0.8614	0.9475	0.9418	0.9513	0.9592	0.9583
Blue	SROCC	0.9261	0.9341	0.9303	0.7827	0.9345	0.9283	0.9375	0.9403	0.9332
Diui	KROCC	0.7677	0.7859	0.7859	0.6020	0.7859	0.7737	-	-	0.7859
	RMSE	6.3973	5.4161	5.3582	7.3504	4.6291	4.8641	-	-	4.1371
FF	PLCC	0.8018	0.8600	0.8548	0.5360	0.7086	0.7583	0.8403	0.8594	0.8620
	SROCC	0.7231	0.8042	0.8154	0.4641	0.4709	0.6882	0.7814	0.7963	0.8286
	KROCC	0.5371	0.6188	0.6289	0.3333	0.4719	0.5131	-	-	0.6403
	RMSE	9.8997	8.4537	8.5971	10.4897	8.7671	8.1001	-	-	6.2993



Fig. 12: Scatter plots of objective scores versus subjective data (DMOS). (a) LIVE Database-Phase I. (b) LIVE Database-Phase II.

distortion type. Table IV gives the detailed information of individual distortion. According to the table, the proposed method is either the best or at parity with the best one for each distortion type. In predicting the quality of distortion types JP2K compression, JPEG compression and Fast Fading, the proposed method has the most significant advantage over the others, especially in JPEG compressed stereo images. Such an outstanding improvement further confirms its efficiency in SIQA. However, it is worth noting that all the metrics seem not to be applicable for JPEG compression, even though a remarkable improvement. As for the prediction of distortion Gaussian Blur, the proposed framework almost have the same correlation coefficients as the best. In terms of distortion White Noise, although the prediction results overtake 0.9, the proposed framework is less prominent than the other 3D metrics. Because the power spectral density of white noise is distributed uniformly in the frequency domain, and localized phase and amplitude maps cannot reflect the changes of image quality. However, in most cases, the proposed framework attains relatively competitive prediction. On the whole, the proposed framework can predict the image quality consistently across different types of distortion.

To verify the demonstration mentioned above, we also report the performances on IRCCyN 3D Image Database according to the corresponding settings. As shown in the Table V, the proposed framework using IS saliency detection model achieves the most competitive performance. The results further confirm the demonstration that cyclopean amplitude and cyclopean phase maps can effectively predict the quality of stereoscopically viewed images under the circumstance of  $\gamma = 0.75$ .

TABLE V: Overall performances on IRCCyN Database

Criteria	PLCC	SROCC	KROCC	RMSE
Benoit [61]	0.4642	0.4558	0.3716	19.5423
Lin [24]	0.6829	0.6196	0.4672	16.1171
Chen [28]	0.6779	0.6376	0.4737	16.2190
Proposed-GBVS	0.6764	0.6058	0.4557	16.2494
Proposed-MSS	0.6107	0.5403	0.3938	17.4700
Proposed-LFWT	0.6822	0.6184	0.4672	16.1313
Proposed-IS	0.7079	0.6435	0.4806	15.5825

To summarize, SIQA is a complicated process, not only distortion-relevant image features but also binocular visual properties should be taken into consideration. The proposed method takes advantages of the two determinant factors to predict perceived quality from the aspects of cyclopean amplitude and phase maps which innovatively combine image features together with binocular visual properties. And the superiority has been further confirmed by experiments.

### V. CONCLUSIONS

In this paper, a full reference quality assessment framework based on cyclopean amplitude and cyclopean phase is proposed for stereo images, in which binocular visual properties and image low-level features are taken into consideration. Based on the fact that amplitude places an emphasis on image degradation, we made further modification on amplitude with the help of saliency detection models. At the same time, by fully considering the binocular rivalry properties, we create the cyclopean amplitude map and the cyclopean phase map. With this method, image features and visual properties are ingeniously integrated with each other. Experiments further confirm the demonstration that cyclopean amplitude and cyclopean phase maps take effect in image quality determination.

In the future, we will pay much attention to the research on binocular visual properties, and explore more effective ways to describe image degradation. At the same time, predicting the quality of asymmetrically distorted stereo images would be another focus.

#### REFERENCES

- K. Lee and S. Lee, "3d perception based quality pooling: Stereopsis, binocular rivalry, and binocular suppression," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 9, no. 3, pp. 533–545, 2015.
- [2] Z. Liu, W. Zou, and O. Le Meur, "Saliency tree: A novel saliency detection framework," *Image Processing, IEEE Transactions on*, vol. 23, no. 5, pp. 1937–1952, 2013.
- [3] L. Zhang, Y. Shen, and H. Li, "Vsi: A visual saliency induced index for perceptual image quality assessment," *Image Processing, IEEE Transactions on*, vol. 23, no. 10, 2014.
- [4] A. K. Moorthy and A. C. Bovik, "Visual importance pooling for image quality assessment," *Selected Topics in Signal Processing, IEEE Journal* of, vol. 3, no. 2, pp. 193–201, 2009.

IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING

- [5] S. Winkler, "Analysis of public image and video databases for quality assessment," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 6, no. 6, pp. 616–625, 2012.
- [6] Y. J. Jung, H. Sohn, S.-I. Lee, H. W. Park, and Y. M. Ro, "Predicting visual discomfort of stereoscopic images using human attention model," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 23, no. 12, pp. 2077–2082, 2013.
- [7] P. Lebreton, A. Raake, M. Barkowsky, and P. Le Callet, "Open perceptual binocular and monocular descriptors for stereoscopic 3d images and video characterization," in *Quality of Multimedia Experience (QoMEX)*, 2015 Seventh International Workshop on. IEEE, 2015, pp. 1–6.
- [8] S. Ryu and K. Sohn, "No-reference quality assessment for stereoscopic images based on binocular quality perception," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 24, no. 4, pp. 591–602, 2014.
- [9] L. Ma, W. Lin, C. Deng, and K. N. Ngan, "Image retargeting quality assessment: a study of subjective scores and objective metrics," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 6, no. 6, pp. 626–639, 2012.
- [10] Z. Wang, E. P. Simoncelli, and 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, vol. 2. Ieee, 2003, pp. 1398–1402.
- [11] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *Image Processing, IEEE Transactions on*, vol. 15, no. 2, pp. 430–444, 2006.
- [12] F. Gao, D. Tao, X. Gao, and X. Li, "Learning to rank for blind image quality assessment," *Neural Networks and Learning Systems, IEEE Transactions on*, vol. 26, no. 10, pp. 2275–2290, 2015.
- [13] W. Hou, X. Gao, D. Tao, and X. Li, "Blind image quality assessment via deep learning," *Neural Networks and Learning Systems, IEEE Transactions on*, vol. 26, no. 6, pp. 1275–1286, 2015.
- [14] P. Ye, J. Kumar, L. Kang, and D. Doermann, "Unsupervised feature learning framework for no-reference image quality assessment," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on.* IEEE, 2012, pp. 1098–1105.
- [15] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *Image Processing, IEEE Transactions* on, vol. 24, no. 8, pp. 2579–2591, 2015.
- [16] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the dct domain," *Image Processing, IEEE Transactions on*, vol. 21, no. 8, pp. 3339–3352, 2012.
- [17] C.-C. Su, L. K. Cormack, and A. C. Bovik, "Closed-form correlation model of oriented bandpass natural images," *Signal Processing Letters, IEEE*, vol. 22, no. 1, pp. 21–25, 2015.
- [18] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *Signal Processing Letters, IEEE*, vol. 20, no. 3, pp. 209–212, 2013.
- [19] J. You, L. Xing, A. Perkis, and 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.
- [20] E. Bosc, R. Pepion, P. Le Callet, M. Koppel, P. Ndjiki-Nya, M. Pressigout, and L. Morin, "Towards a new quality metric for 3-d synthesized view assessment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 7, pp. 1332–1343, 2011.
- [21] C. Fehn, "Depth-image-based rendering (dibr), compression, and transmission for a new approach on 3d-ty," in *Electronic Imaging 2004*. International Society for Optics and Photonics, 2004, pp. 93–104.
- [22] F. Shao, W. Lin, S. Gu, G. Jiang, and T. Srikanthan, "Perceptual fullreference quality assessment of stereoscopic images by considering binocular visual characteristics." *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*, vol. 22, no. 5, pp. 1940–1953, 2013.
- [23] S. Ryu, D. H. Kim, and K. Sohn, "Stereoscopic image quality metric based on binocular perception model," in *Image Processing (ICIP), 2012* 19th IEEE International Conference on. IEEE, 2012, pp. 609–612.
- [24] Y.-H. Lin and J.-L. Wu, "Quality assessment of stereoscopic 3d image compression by binocular integration behaviors," *Image Processing*, *IEEE Transactions on*, vol. 23, no. 4, pp. 1527–1542, 2014.
- [25] X. Wang, S. Kwong, and Y. Zhang, "Considering binocular spatial sensitivity in stereoscopic image quality assessment," in *Visual Communications and Image Processing (VCIP), 2011 IEEE*. IEEE, 2011, pp. 1–4.
- [26] W. J. Levelt, On binocular rivalry. Mouton The Hague, 1968, vol. 2.

- [27] A. Maalouf and M.-C. Larabi, "Cyclop: A stereo color image quality assessment metric," in Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on. IEEE, 2011, pp. 1161–1164.
- [28] M.-J. Chen, C.-C. Su, D.-K. Kwon, L. K. Cormack, and A. C. Bovik, "Full-reference quality assessment of stereopairs accounting for rivalry," *Signal Processing: Image Communication*, vol. 28, no. 9, pp. 1143– 1155, 2013.
- [29] F. Shao, K. Li, W. Lin, G. Jiang, M. Yu, and Q. Dai, "Full-reference quality assessment of stereoscopic images by learning binocular receptive field properties," *Image Processing, IEEE Transactions on*, vol. 24, no. 10, pp. 2971–2983, 2015.
- [30] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 207, no. 1167, pp. 187–217, 1980.
- [31] M. C. Morrone and D. Burr, "Feature detection in human vision: A phase-dependent energy model," *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 235, no. 1280, pp. 221–245, 1988.
- [32] P. Kovesi, "Image features from phase congruency," Videre: Journal of computer vision research, vol. 1, no. 3, pp. 1–26, 1999.
- [33] Z. Liu and R. Laganière, "Phase congruence measurement for image similarity assessment," *Pattern Recognition Letters*, vol. 28, no. 1, pp. 166–172, 2007.
- [34] Z. Wang and E. P. Simoncelli, "Local phase coherence and the perception of blur," in Advances in neural information processing systems, 2003, p. None.
- [35] R. Hassen, Z. Wang, and M. Salama, "No-reference image sharpness assessment based on local phase coherence measurement," in *Acoustics Speech and Signal Processing (ICASSP)*, 2010 IEEE International Conference on. IEEE, 2010, pp. 2434–2437.
- [36] G. Blanchet, L. Moisan, and B. Rougé, "Measuring the global phase coherence of an image," in *Image Processing*, 2008. *ICIP* 2008. 15th IEEE International Conference on. IEEE, 2008, pp. 1176–1179.
- [37] M. Narwaria, W. Lin, I. V. McLoughlin, S. Emmanuel, and L.-T. Chia, "Fourier transform-based scalable image quality measure," *Image Processing, IEEE Transactions on*, vol. 21, no. 8, pp. 3364–3377, 2012.
- [38] M. von Grunau and M. Iordanova, "Visual selection: Facilitation due to stimulus saliency," in *Cybernetic Vision*, 1996. Proceedings., Second Workshop on. IEEE, 1996, pp. 15–20.
- [39] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in Advances in neural information processing systems, 2006, pp. 545–552.
- [40] R. Achanta and S. Süsstrunk, "Saliency detection using maximum symmetric surround," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010, pp. 2653–2656.
- [41] N. İmamoğlu, W. Lin, and Y. Fang, "A saliency detection model using low-level features based on wavelet transform," *Multimedia, IEEE Transactions on*, vol. 15, no. 1, pp. 96–105, 2013.
- [42] X. Hou, J. Harel, and C. Koch, "Image signature: Highlighting sparse salient regions," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 1, pp. 194–201, 2012.
- [43] K. Gu, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, "Automatic contrast enhancement technology with saliency preservation," *Circuits* and Systems for Video Technology, IEEE Transactions on, vol. 25, no. 9, pp. 1480–1494, 2015.
- [44] Y. Fang, J. Wang, M. Narwaria, P. Le Callet, and W. Lin, "Saliency detection for stereoscopic images," *Image Processing, IEEE Transactions* on, vol. 23, no. 6, pp. 2625–2636, 2014.
- [45] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "Live image quality assessment database release 2," 2005.
- [46] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and M. Gross, "Nonlinear disparity mapping for stereoscopic 3d," ACM Transactions on Graphics (TOG), vol. 29, no. 4, p. 75, 2010.
- [47] L. Jansen, S. Onat, and P. König, "Influence of disparity on fixation and saccades in free viewing of natural scenes," *Journal of Vision*, vol. 9, no. 1, p. 29, 2009.
- [48] X. K. Yang, W. S. Lin, Z. Lu, E. P. Ong, and S. Yao, "Just-noticeabledistortion profile with nonlinear additivity model for perceptual masking in color images," in *IEEE International Conference on Acoustics*, 2003, pp. III–609–12 vol.3.
- [49] Y. Zhao, Z. Chen, C. Zhu, Y.-P. Tan, and L. Yu, "Binocular justnoticeable-difference model for stereoscopic images," *Signal Processing Letters, IEEE*, vol. 18, no. 1, pp. 19–22, 2011.
- [50] W. Zhou, G. Jiang, M. Yu, F. Shao, and Z. Peng, "Pmfs: A perceptual modulated feature similarity metric for stereoscopic image quality assessment," *Signal Processing Letters, IEEE*, vol. 21, no. 8, pp. 1003– 1006, 2014.

- [51] N. Qian and S. Mikaelian, "Relationship between phase and energy methods for disparity computation," *Neural Computation*, vol. 12, no. 2, pp. 279–292, 2000.
- [52] M. C. Morrone and R. A. Owens, "Feature detection from local energy," *Pattern recognition letters*, vol. 6, no. 5, pp. 303–313, 1987.
- [53] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *JOSA A*, vol. 4, no. 12, pp. 2379– 2394, 1987.
- [54] L. Zhang, D. Zhang, and X. Mou, "Fsim: a feature similarity index for image quality assessment," *Image Processing, IEEE Transactions on*, vol. 20, no. 8, pp. 2378–2386, 2011.
- [55] F. Qi, T. Jiang, S. Ma, and D. Zhao, "Quality of experience assessment for stereoscopic images," in *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on.* IEEE, 2012, pp. 1712–1715.
- [56] C.-C. Su, A. C. Bovik, and 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.
- [57] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.
- [58] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images," *EURASIP journal on image and video processing*, 2008.
- [59] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment, phase ii," 2003.
- [60] S. Li, F. Zhang, L. Ma, and K. N. Ngan, "Image quality assessment by separately evaluating detail losses and additive impairments," *Multimedia, IEEE Transactions on*, vol. 13, no. 5, pp. 935–949, 2011.
- [61] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Using disparity for quality assessment of stereoscopic images," in *Image Processing*, 2008. ICIP 2008. 15th IEEE International Conference on. IEEE, 2008, pp. 389–392.



Wen Lu received the M.S and Ph.D. degrees in pattern recognition and intelligent systems at Xidian University, in 2006 and 2009, respectively. He is an associate professor at Xidian University. His research interests include image and video understanding, visual quality assessment, and computational vision.



Qinggang Meng received B.S. and M.S. degrees in School of Electronic Information Engineering, Tianjin University, China, and Ph.D. degree in Computer Science from Aberystwyth University, UK. He is a Senior Lecturer in the Department of Computer Science, Loughborough University, UK. His research interests include pattern recognition, artificial intelligence and computer vision. He is a senior member of the IEEE and a fellow of the Higher Education Academy of United Kingdom.



**Yancong Lin** is a M.S. student with School of Electronic Information Engineering, Tianjin University, Tianjin, China. His research interests include stereo vision research, stereo image displaying and visual quality assessment.



Zhihan Lv was granted PhD. degree in computer applied technology from Ocean university of China (2006-2012). Before that, he has a full-time research experience at Centre National de la Recherche Scientifique (CNRS)-UPR9080 in Paris (2010-2011). After then, he has fulfilled a two-year postdoc research experience at Umea university and an invited teaching experience at KTH Royal Institute of Technology in Sweden. Since 2012, he has held an assistant professor position at Chinese Academy of Science. His research interests include computer vision, 3D

visualization, and human-computer interaction.



Jiachen Yang received the M.S. and Ph.D. degrees in Communication and Information Engineering from the Tianjin University, Tianjin, China, in 2005 and 2009, respectively. He is an associate professor at Tianjin University. In 2014, he was a visiting scholar in the Department of Computer Science, School of Science at Loughborough University, UK. His research interests include stereo camera, stereo vision research, pattern recognition, stereo image displaying and quality evaluation.



Houbing Song received the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, VA, in 2012. In 2012, he joined the Department of Electrical and Computer Engineering, West Virginia University, Montgomery, WV, where he is currently an Assistant Professor and the founding director of both the West Virginia Center of Excellence for Cyber-Physical Systems (WVCECPS) sponsored by the West Virginia Higher Education Policy Commission and the Security and Optimization for Networked Globe Laboratory

(SONG Lab). His research interests lie in the areas of cyber-physical systems, internet of things and digital image processing. Dr. Song's research has been supported by the West Virginia Higher Education Policy Commission. Dr. Song is a senior member of IEEE and a member of ACM.