

# Blind Quality Metric of DIBR-Synthesized Images in the Discrete Wavelet Transform Domain

Guangcheng Wang<sup>ID</sup>, Zhongyuan Wang<sup>ID</sup>, Member, IEEE, Ke Gu<sup>ID</sup>, Member, IEEE, Leida Li<sup>ID</sup>,  
Zhifang Xia, and Lifang Wu<sup>ID</sup>

**Abstract**—Free viewpoint video (FVV) has received considerable attention owing to its widespread applications in several areas such as immersive entertainment, remote surveillance and distanced education. Since FVV images are synthesized via a depth image-based rendering (DIBR) procedure in the “blind” environment (without reference images), a real-time and reliable blind quality assessment metric is urgently required. However, the existing image quality assessment metrics are insensitive to the geometric distortions engendered by DIBR. In this research, a novel blind method of DIBR-synthesized images is proposed based on measuring geometric distortion, global sharpness and image complexity. First, a DIBR-synthesized image is decomposed into wavelet subbands by using discrete wavelet transform. Then, the Canny operator is employed to detect the edges of the binarized low-frequency subband and high-frequency subbands. The edge similarities between the binarized low-frequency subband and high-frequency subbands are further computed to quantify geometric distortions in DIBR-synthesized images. Second, the log-energies of wavelet subbands are calculated to evaluate global sharpness in DIBR-synthesized images. Third, a hybrid filter combining the autoregressive and bilateral filters is adopted to compute image complexity. Finally, the overall quality score is derived to normalize geometric distortion and global sharpness by the image complexity. Experiments show that our proposed quality method is superior to the competing reference-free state-of-the-art DIBR-synthesized image quality models.

**Index Terms**—Depth image-based rendering, blind quality assessment, geometric distortion, global sharpness, image complexity.

Manuscript received December 5, 2018; revised August 11, 2019; accepted September 19, 2019. Date of publication October 10, 2019; date of current version November 27, 2019. This work was supported in part by the National Key Research and Development Project under Grant 2016YFE0202300, in part by the National Natural Science Foundation of China under Grant 61671332, Grant U1736206, Grant 61703009, Grant 61771473, and Grant 61379143, in part by the Hubei Province Technological Innovation Major Project under Grant 2019AAA049, and in part by the Natural Science Fund of Hubei Province under Grant 2018CFA024. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Laura Toni. (*Corresponding authors:* Zhongyuan Wang; Ke Gu.)

G. Wang and Z. Wang are with the National Engineering Research Center for Multimedia Software, School of Computer Science, Wuhan University, Wuhan 430072, China (e-mail: wangguangcheng0428@163.com; wzy\_hope@163.com).

K. Gu is with the Beijing Key Laboratory of Computational Intelligence and Intelligent System, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China (e-mail: guke.doctor@gmail.com).

L. Li is with the School of Artificial Intelligence, Xidian University, Xian 710071, China (e-mail: ldli@xidian.edu.cn).

Z. Xia is with the Beijing Advanced Innovation Center for Future Internet Technology, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China, and also with the State Information Center, Beijing 100045, China (e-mail: spidergirl21@163.com).

L. Wu is with the College of Information and Communication Engineering, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China (e-mail: lfwu@bjut.edu.cn).

Digital Object Identifier 10.1109/TIP.2019.2945675

## I. INTRODUCTION

WITH the rapid development of 3D-related technologies, many new challenges have emerged. Free viewpoint video (FVV), due to its comprehensive applications in immersive entertainment, remote surveillance and distanced education, has received extensive attentions and been regarded as a new important direction of video technology development. FVV promotes the rapid development of Depth Image-Based Rendering (DIBR) technology, producing new viewpoints from multiple views. Unfortunately, the major problem of the DIBR algorithms is the disocclusion: regions which are occluded in the captured views become visible in the virtual ones. Due to the lack of original texture information, a synthesized image often contains disocclusion holes which significantly degrades the quality [1], [2]. The disoccluded regions contained in the synthesized views usually produce some geometric distortion, as shown in Fig. (b)-(d), which has many different structural characteristics as compared with the distortions in natural scene images [1]. However, the existing Image Quality Assessment (IQA) methods, mainly devised for natural scene images, work poorly in capturing the geometric distortions that occur in DIBR-synthesized images. With such consideration, it is necessary to design a reliable objective assessment metric for predicting the quality of DIBR-synthesized images.

In the past, IQA has attracted a wide range of attentions from the relevant researchers. IQA can be divided into subjective and objective assessment [3]. Subjective assessment means the tester evaluates the quality of the images and videos through some equipments. In general, the Human Visual System (HVS) is the final receiver of visual information such as images and videos. So subjective assessment is thought of the most reliable evaluation method, which is usually used to produce ground truth for objective assessment. Subjective assessment is expensive and time-consuming, so it is difficult to apply in large-scale practical applications. Therefore, extensive researches have been devoted to objective quality assessment, which adopts mathematical models to simulate the HVS, so objective assessment can be used to effectively evaluate the quality of visual information and save human resources. According to whether the reference information is needed, objective IQA metrics can be divided into reference-based and reference-free/blind metrics. For instance, Wang *et al.* devised a reference-based IQA metric, namely Structural SIMilarity (SSIM) [4], which is based upon the differences of contrast, luminance and structural similarity between a distorted image

and its associated reference image. Gu *et al.* proposed a reference-free quality metric with the free energy principle and important HVS inspired features [5]. Moreover, the objective evaluation models, e.g. PSNR and SSIM, are widely used to evaluate visual tasks, such as image fusion and image enhancement. [6]–[11].

Generally, according to the types of distortions, objective IQA metrics can further be divided into two classes. The first class is devised for evaluating specific distortions, which include noise, blockiness, sharpness and contrast distortions, etc. Such as, Vu *et al.* proposed the algorithm S<sub>3</sub>, which is designed based on both spectral and spatial properties to measure local perceived sharpness [13]. Gu *et al.* designed a reference-free quality metric based on analyzing the parameters of classical autoregressive (AR) [14]. Another class is the general-purpose algorithms, which can evaluate multiple distortions in distorted images. Most of the existing general-purposed metrics were proposed based on feature aggregation and natural scene statistics (NSS) [15]–[20]. Usually, the process of feature aggregation is composed of two steps: 1) extracting the perceptual features (e.g. texture, luminance, contrast and color) of the distorted images; 2) using machine learning tools to learn these features to obtain a quality evaluator.

Since the existing IQA metrics, mainly devised for natural scene images, have difficulty in capturing the geometric distortions in DIBR-synthesized images, several recent studies have been proposed to evaluate DIBR-synthesized images. Battisti *et al.* proposed 3D-SWIM model [21], which analyzes the similarities of statistical features from wavelet subbands between the reference and distorted DIBR-synthesized images. Sandić-Stanković *et al.* presented MW-PSNR [22] and MP-PSNR [23] based on the morphological wavelet decomposition and morphological pyramid decomposition, respectively. In order to improve the performance of MP-PSNR, Sandić-Stanković *et al.* further put forward MP-PSNR-reduce, which calculates squared error maps between the appropriate detail images only at higher pyramids' scales of the two pyramids [24]. The MP-PSNR-reduce has lower computational complexity than MP-PSNR. Jakhetiya *et al.* designed a reference-based IQA metric for DIBR-synthesized images based on perceptual relevant prediction model, which emphasizes the importance of textual regions in visual quality assessment [25]. Li *et al.* designed a reference-based IQA algorithm measuring LOcal Geometric distortions in disoccluded regions and global Sharpness (LOGS) caused by warping and rendering [26]. Farid *et al.* proposed a block-based quality assessment algorithm to extract the phase congruency based perceptual features from the input views (left and right view images) and the virtual image, and compare the obtained perceptual features to predict the quality of the DIBR-synthesized views [27]. Ling *et al.* predicted the quality of DIBR-synthesized images by taking advantage of the characteristics of convolutional sparse coding, which can be used to learn from local regions and generate one sparse representation to learn the impact of non-uniform artifacts on visual quality [28]. Ling *et al.* proposed a Sketch-Token-based video quality assessment metric checking how the classes

of contours change between the reference and the degraded sequences spatially and temporally [29]. Assuming that the high-quality and DIBR-synthesized images possess different responses on morphological operations, Tian *et al.* [30] proposed a morphological based approach (named NIQSV) to detect the geometric distortion and consequently measure the image quality. Later, they extended NIQSV (to NIQSV+) to further improve the performance by integrating scores of NIQSV, stretching and black hole together [31]. In [32], a new NSS model based on AR-based local image description was proposed by Gu *et al.*, namely AR-plus thresholding, which can evaluate the DIBR-synthesized views effectively. Yue *et al.* introduced a DIBR-Synthesized IQA metric based upon Combining Local and Global Measures (CLGM) [33]. Jakhetiya *et al.* assumed geometric distortions behave like outliers and empirically proved this assumption using the three *sigma* rule-based robust outlyingness ratio statistics. With this view, the authors proposed an IQA algorithm using median filtering to highlight both geometric and structural distortions and further fused these highlighted distortions to predict the quality of DIBR-synthesized images [34]. Zhou *et al.* presented a reference-free quality index for synthesized views using the difference-of-gaussian representation based edge statistics and texture naturalness [35].

From the existing IQA models of DIBR-synthesized images, we find two major deficiencies. On the one hand, the reference-based IQA modules designed for DIBR-synthesized images are all based on the reference images, but FVV images are synthesized by DIBR in “blind” environment. Reference-free IQA algorithms are more valuable than the reference-based IQA methods. On the other hand, the aforementioned reference-free IQA algorithms are deficient in performance (e.g. accuracy and monotonicity separately measured by PLCC and SRCC as shown in experimental results such as Table I) or computational efficiency (seeing Table IV). Thus, a reliable and real-time blind IQA metric of DIBR-synthesized images is particularly favorite. As shown in Fig.1 (b)-(d) and Fig.1 (f), the geometric distortion and blurring effect seriously destroy the visual quality of DIBR-synthesized views. Therefore, in order to effectively evaluate the DIBR-synthesized views, a novel reference-free quality assessment method is proposed based on measuring the geometric distortion and global sharpness. Moreover, the image complexity is employed to weaken the impact of image content diversity on the proposed reference-free IQA model. A DIBR-synthesized image is first decomposed into wavelet subbands (LL, LH, HL and HH subbands) by using the Cohen-Daubechies-Feauveau 9/7 filter [36]. Then, we adopt the Canny operator [38] to extract the edges of different wavelet subbands. We can only detect the edges of geometric distortion in the binarized low-frequency subband (LL subband) because the high-frequency information of DIBR-synthesized images is separated from the low-frequency subband. In contrast, the edges detected in high-frequency subbands (LH, HL and HH subbands) consist of the edges of DIBR-synthesized views themselves and the geometric distortions contained in these views as well. The edge similarities between the low-frequency subband and the high-frequency subbands can describe the degree of geometric

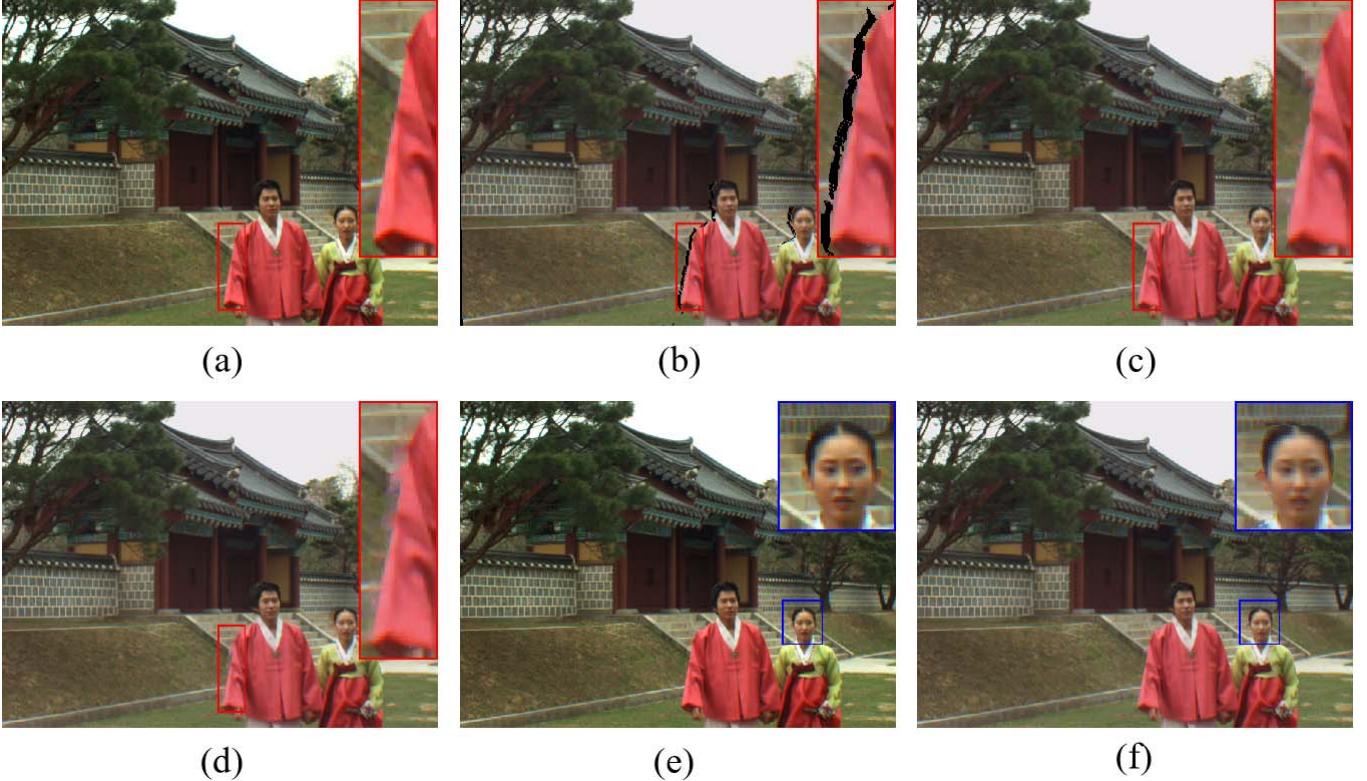


Fig. 1. An example of a reference image and its associated distorted DIBR-synthesized views. (a) and (e) are reference images, (b)-(d) and (f) are distorted DIBR-synthesized views, (b)-(d) contain obvious geometric distortions compared to the reference image, (e) includes obvious blur distortion compared to the reference image.

distortion in DIBR-synthesized views. Moreover, because the blurring mainly occurs in the high-frequency component of the blurred images, we use the log-energies of low- and high-frequency subbands to evaluate the global sharpness in DIBR-synthesized images and give higher weights to the log-energies of high-frequency subbands. Finally, we adopt a hybrid filter combining the AR and bilateral (BL) filters to compute image complexity, which is used to normalize geometric distortion and global sharpness to achieve the overall quality score. AR and BL are respectively used to describe the texture and edge regions of DIBR-synthesized images. We verify the performance of the proposed algorithm and state-of-the-art IQA methods on two public datasets, IRCCyN/IVC [1] and IETR [2] DIBR databases, which focus on the distortions only caused by DIBR algorithms and are often employed to verify the performance of DIBR-synthesized image quality assessment models. As one widely accepted data representation format for 3D scenes, the Multiview Video plus Depth (MVD) format [39] consists of multiple texture images and their corresponding depth maps at some particular viewpoints. The MVD test sequences used in IRCCyN/IVC and IETR databases are provided by MPEG for the 3D video coding. The IRCCyN/IVC and IETR databases both include seven DIBR algorithms and invite more than 40 naive (non-expert) observers to conduct subjective experiments. Compared with the IRCCyN/IVC database, the DIBR algorithms used in the IETR dataset have better performance, and so the synthesized views in IETR database contain less geometric distortions

(eg. Black holes). A larger distance between reference cameras and “virtual” camera leads to larger disoccluded regions. Since the proposed algorithm is only based on measuring the geometric distortion in disoccluded regions and global sharpness of the synthesized views, the proposed algorithm does not depend on the type of DIBR algorithms and the distance between reference cameras and “virtual” camera.

The rest of the paper is outlined as follows. Section II presents the geometric distortion, global sharpness and image complexity models and the final blind IQA metric. Section III conducts comparisons of our reference-free metric with state-of-the-art reference-based and reference-free IQA methods devised for natural scene images and DIBR-synthesized views. Section IV concludes the whole paper.

## II. PROPOSED QUALITY ASSESSMENT METRIC

Fig. 1 shows an example of a reference image and its associated DIBR-synthesized image. Apparently, as compared with the reference image, the quality of a DIBR-synthesized image is highly degraded by the geometric distortion and global sharpness. In Fig. 1(b)-(d), geometric distortions marked by red boxes have a remarkable destructive effect on naturalness attribute of a DIBR-synthesized image. From Fig. 1(f), we can observe the blur distortion in the distorted DIBR-synthesized image. In this research, we measure the geometric distortion and global sharpness of DIBR-synthesized images in the discrete wavelet transform domain. In order to further weaken the impact of image content diversity on our proposed IQA model,

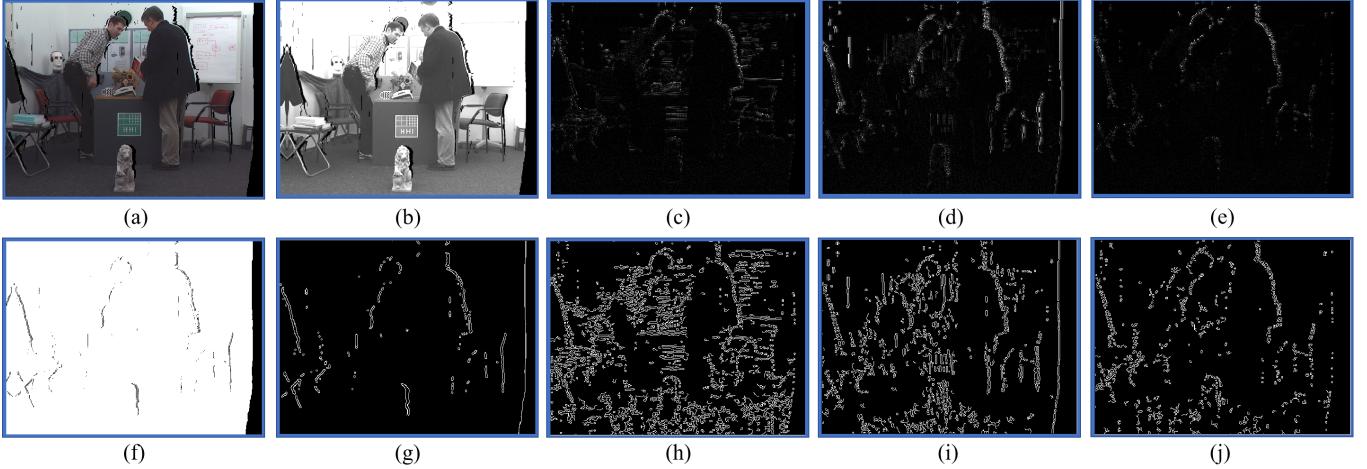


Fig. 2. An example of a DIBR-synthesized image and its associated wavelet subbands and the edge detection of wavelet subbands. (a) a DIBR-synthesized image, (b)-(e) are the LL, HL, LH and HH subbands of (a) respectively, (f) is the binarized LL subband, (g)-(j) are the edge detection results of (f) and (c)-(e).

image complexity is computed to normalize the geometric distortion and global sharpness.

#### A. Geometric Distortion Detection and Quantization

The key problem in effectively evaluating DIBR-synthesized images is whether the metric can accurately detect and quantify geometric distortion. In this research, the DIBR-synthesized image is first decomposed into low- and high-frequency subbands by adopting the Cohen-Daubechies-Feauveau 9/7 filter [36]. Fig. 2 (a)-(e) show an example of a DIBR-synthesized image and its corresponding wavelet subbands. In order to accurately extract the edges of the geometric distortion from LL subband, we further binarize the LL subband through a threshold computed by using a graythresh function [37], as shown in Fig. 2 (f). Then, the Canny operator [38] is used to detect the edges of the binarized LL (BLL) subband and high-frequency wavelet subbands as shown below.

$$C_M = \text{Canny}(M), \quad (1)$$

where  $M$  represents the BLL, HL, LH and HH subbands.

Fig. 2 (g)-(j) show the edge detection results of BLL, HL, LH and HH subband from a given DIBR-synthesized view, respectively. From Fig. 2 (g)-(j), we can see that the edges of BLL subband are only the edges of geometric distortions, while the edges of HL, LH and HH subbands are the mixed edges of the DIBR-synthesized view itself and its geometric distortions. The edge similarities between BLL and high-frequency subbands can represent the degree of geometric distortion in the DIBR-synthesized view. Therefore, the structural similarities between  $C_{BLL}$  and  $C_{HL}$ ,  $C_{LH}$  and  $C_{HH}$  are computed to quantify geometric distortion. The method of calculation is provided as follows:

$$S_H = \frac{1}{L} \sum_{l=1}^L \left( \frac{2C_{BLL}(l) \cdot C_{HL}(l) + \varepsilon}{C_{BLL}(l) + C_{HL}(l) + \varepsilon} \right), \quad (2)$$

$$S_V = \frac{1}{L} \sum_{l=1}^L \left( \frac{2C_{BLL}(l) \cdot C_{LH}(l) + \varepsilon}{C_{BLL}(l) + C_{LH}(l) + \varepsilon} \right), \quad (3)$$

$$S_D = \frac{1}{L} \sum_{l=1}^L \left( \frac{2C_{BLL}(l) \cdot C_{HH}(l) + \varepsilon}{C_{BLL}(l) + C_{HH}(l) + \varepsilon} \right), \quad (4)$$

where  $\varepsilon$  is a constant assigned as one for avoiding the problem of zero denominator;  $L$  represents the number of the pixels in image;  $l$  is the pixel index. We quantify geometric distortion from horizontal, vertical and diagonal directions. This is also closer to the multi-directional character of the HVS. The final quantization result of geometric distortion in DIBR-synthesized images is as follows:

$$Q_1 = \sum \omega_i \cdot S_i, \quad (5)$$

where  $i = H, V$  and  $D$ ; The weight coefficients  $\omega_i$  are used to adjust the proportions of the  $S_H$ ,  $S_V$  and  $S_D$ ; Here,  $\omega_i$  are simply set to 1 in this work. Future work can be devoted to finding more suitable parameters for possibly further improving the performance of the proposed model.

#### B. Global Sharpness Evaluation

In addition to the geometric distortion in DIBR-synthesized images, the warping and rendering in DIBR also produce blurring effect, which mainly exhibits around the transitions of background and foreground. Therefore, in the process of evaluating DIBR-synthesized images, we should take the primary factor of sharpness into consideration. Since blurring effect mainly occurred in the high-frequency part of the blurred images, many IQA metrics designed for sharpness have been proposed based upon analyzing the high-frequency information of blurred images [40]–[43]. Inspired the research [44], we propose a simple and reliable metric to evaluate the global sharpness. Different from the literature [44], we consider that the influence of low-frequency information on the engendering of blur distortion cannot be completely ignored. Thus, we assign different weights to the log-energies of the low- and high-frequency subbands and the proportion of high-frequency subbands is greater than low-frequency subband.

The method for global sharpness evaluation is composed of two parts. First, we calculate the log-energy at each wavelet

subband, which can be divided into two steps: 1) calculating the energy of the wavelet subband image (i.e., the square of the pixel values); 2) computing the logarithm of the energy sum of the whole image [44]. The calculation method is defined as follows:

$$E_{XY} = \log_{10}(1 + \frac{1}{L} \sum_{l=1}^L XY^2(l)), \quad (6)$$

where  $XY$  represents the LL, LH, HL or HH subbands;  $l$  is the pixel index;  $L$  represents the number of the wavelet coefficients in each subband.

Next, we measure the total log-energies at all decomposed wavelet subbands via

$$Q_2 = a \cdot E_{HH} + b \cdot \frac{E_{HL} + E_{LH}}{2} + c \cdot E_{LL}. \quad (7)$$

Because the blurring effects mainly occur in high-information of the DIBR-synthesized images and HH subband can be regarded to span a higher radial spatial frequency (by a factor of  $\sqrt{2}$ ) than the LH and HL bands, the parameters  $a = 0.5$ ,  $b = 0.3$  and  $c = 0.2$  are chosen to weight the log-energies of different wavelet subbands. The obtained  $Q_2$  is used to evaluate the global sharpness in DIBR-synthesized images.

### C. Image Complexity Estimation

Image complexity is an essential concept in human basic perception to visual stimulus. Image complexity is a key factor to be considered when designing DIBR-synthesized IQA metric, since it relates to the effects of gaze direction and spatial masking. In general, high-complexity images contain more high-frequency information, including edges and textures. It is obvious that low-complexity has high self-description ability compared to high-complexity, since textures and edges are more difficult to self-described than smooth regions. A hybrid filter systematically combining the AR and BL filters has been used in the past to estimate image complexity [45]–[47]. The expression of the AR filter is as follows:

$$y_q = \Phi^n(x_q)\mathbf{a} + b_i, \quad (8)$$

where  $x_q$  is the value of the pixel located at  $i$  in the given image;  $\Phi^n(x_q)$  contains the  $n$  neighboring pixels of  $x_q$ ;  $\mathbf{a} = (a_1, a_2, \dots, a_n)^T$  is a vector of AR parameters;  $b_i$  is the residual error between a given pixel and its corresponding predicted result by AR filtering. Then the predicted image is

$$\hat{y}_q = \Phi^n(x_q)\hat{\mathbf{a}}, \quad (9)$$

where  $\hat{\mathbf{a}}$  is calculated based upon the method in [48].

Compared with texture regions, the performance of AR operator in edge regions is unsatisfied. In order to overcome this shortcoming, the classical BL filter is employed to combine AR operator towards a tradeoff better filter. The BL filter has edge-preserving function, as defined as follows:

$$y_q = \Phi^n(x_q)\mathbf{b} + d_i, \quad (10)$$

where  $\mathbf{b} = (b_1, b_2, \dots, b_n)^T$  are a set of coefficients produced by BL filtering;  $d_i$  is the error between a given pixel and its associated output prediction. The parameters of BL

filter follow the assignment in [49]. The BL filter provides better results around the luminance edges than the AR-based predictor, but it has limitation in maintaining texture details. A hybrid filter was designed by combining the AR and BL filters to obtain the best performance for edge and texture regions. The hybrid filter is defined by

$$\hat{y}_q = \frac{\Phi^n(x_q)\hat{\mathbf{a}} + k\Phi^n(x_q)\mathbf{b}}{1+k}, \quad (11)$$

where  $k = 9$  adjusts the relative strength of the responses of the AR and BL filters. The detailed solution of the coefficients  $k$  can be seen in [46].

The image complexity of a DIBR-synthesized image is estimated as follows:

$$Q_3 = - \int H'_{(\rho)} \log H'_{(\rho)} d\rho, \quad (12)$$

where  $H'_{(\rho)}$  represents the probability density of grayscale  $\rho$  in the error map between the given DIBR-synthesized image and its corresponding filtered result, i.e.,  $\Delta y_q = y_q - \hat{y}_q$ ;  $y_q$  is the value of a pixel at location  $x_q$ .

### D. Proposed DIBR-Synthesized Image Quality Metric

With these quality scores of the geometric distortion ( $Q_1$ ), global sharpness ( $Q_2$ ) and image complexity ( $Q_3$ ), we carefully consider how to effectively integrate these scores to form the final overall quality score.  $Q_1$  is a geometric distortion measure, and a higher  $Q_1$  value indicates the geometric distortion in DIBR-synthesized images is more severe.  $Q_2$  is a global sharpness measure, and a higher  $Q_2$  value denotes that the global sharpness in DIBR-synthesized images is more severe. The image complexity  $Q_3$  is used to normalize the quantized geometric distortion and global sharpness, in which the impact of image content diversity on IQA metric can be eliminated effectively. With these considerations, a simple pooling strategy is proposed to define the overall quality score of a DIBR-synthesized image:

$$Q = \frac{Q_1 + \alpha \cdot Q_2}{1 + \alpha} \cdot \frac{1}{Q_3}, \quad (13)$$

where  $\alpha$  is a normalized parameter for adjusting the contributions of  $Q_1$  and  $Q_2$  to achieve the best performance. In section III, we will explain the selection of the parameter  $\alpha$  in detail. The higher quality DIBR-synthesized images have less geometric distortions and global sharpness. Therefore, the lower  $Q$  value shows the better image quality. We further show an overview of the proposed IQA metric in Fig. 3. In the future, we plan to improve our model based on joint domains and big data analysis like [50] or extend it to video quality assessment by inserting temporal weighting like [51], [52].

## III. EXPERIMENTS

This section concentrates on measuring and comparing the performance of the proposed quality metric with state-of-the-art IQA approaches on the image database specific to DIBR-synthesized IQA.

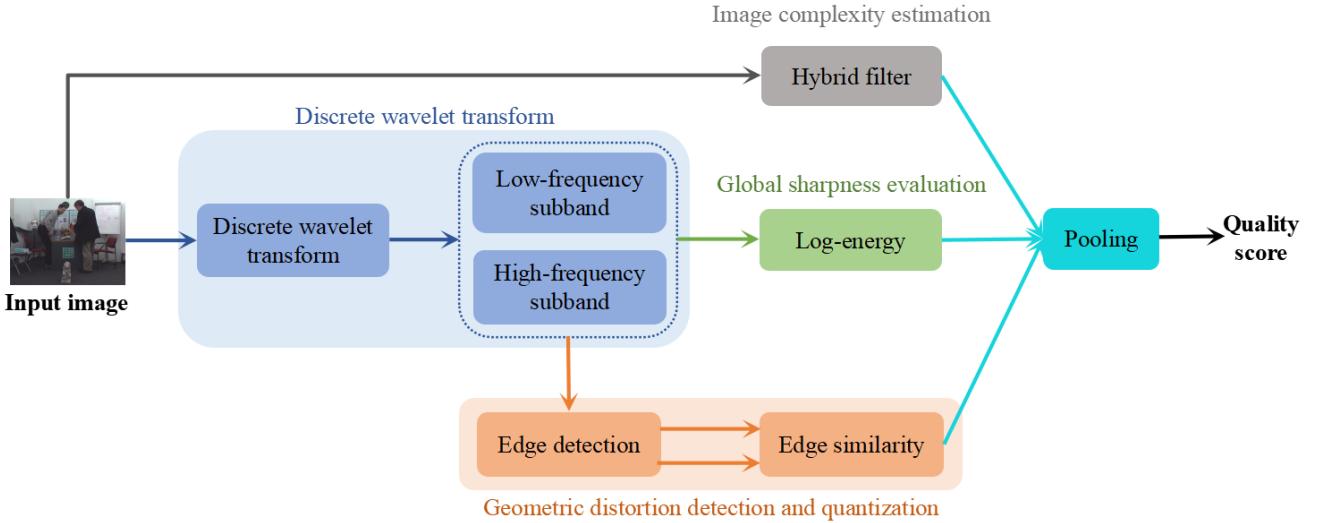


Fig. 3. The overview of the proposed quality assessment metric of DIBR-synthesized images.

#### A. DIBR-Synthesized Image Dataset

1) IRCCyN/IVC Database [1]: The IRCCyN/IVC DIBR-synthesized database is composed of 12 reference images and its associated 84 DIBR-synthesized images produced by 7 different DIBR methods, which are denoted by A1-A6 [53]–[58] and A7 (warping without rendering). The subjective assessment of the IRCCyN/IVC DIBR database [1] is based on the Absolute Category Rating-Hidden Reference (ACR-HR) algorithm [59] and the subjective scores are given in the form of mean opinion score (MOS). Therefore, it is unreasonable to use the MOS values as basic facts for evaluating the objective reference-based methods. Similar to [21], in order to use the subjective scores for reference-based quality assessment, the Difference Mean Opinion Scores (DMOS) are first generated based on the MOS values [60]. The A1-A7 are introduced briefly as follows:

- A1: The depth map is first filtered by a Gaussian low-pass filter. Then, the interpolation algorithm is used to cut the borders and restore the image to its original size [53].
- A2: The depth map is also preprocessed as that in A1 [53]. Then, the borders are dealt with via the inpainted method in [54].
- A3: It is a 3D view generation system proposed by Tanimoto *et al.* [55]. It also uses an inpainting algorithm [54] to fill in the disoccluded region.
- A4: It is a hole filling algorithm embedded in a 3D video system [56]. It adopts the depth information to fill in the disoccluded area.
- A5: It is a hole filling approach for DIBR systems based on patch-based texture synthesis [57]. The unknown areas are synthesized by copying content from the known parts of the image.
- A6: It extends the A5 [57] method by considering the depth temporal information to synthesize the disoccluded region [58].

- A7: It is a method with only a warping step without any postprocessing.

2) IETR Database [2]: The IETR DIBR-synthesized database is composed of 10 reference images and its associated 140 DIBR-synthesized images produced by B1-B7 [62]–[68]. The subjective assessment of the IETR DIBR database [2] is based on the subjective assessment methodology for video quality (SAMVIQ) algorithm [61] and the subjective scores are given in the form of DMOS. The B1-B7 are introduced briefly as follows:

- B1: It employs an exemplar-based texture synthesis technique, which adopts a *confidence* to compute patch priorities and optimize the fill order of the target regions according to their priorities [62].
- B2: It is an object-based layered depth image representation to improve the quality of virtual synthesized views [63].
- B3: It is a depth based disocclusion filling method using patch-based texture synthesis [64].
- B4: It is a hole filling approach for DIBR systems based on background reconstruction, in which the temporal correlation information in both the 2D video and its corresponding depth map are exploited to construct a background video based on motion compensation and modified Gaussian Mixture model [65].
- B5: It uses two pyramid-like approaches, namely Hierarchical Hole-Filling (HHF) and Depth Adaptive Hierarchical Hole-Filling, to eliminate the disoccluded holes in DIBR synthesized views [66].
- B6: The depth discontinuity artifacts are solved by performing a post-filter on the projected depth map. Then, the inpainting method proposed in [54] is used to fill the holes in the disoccluded regions [67].
- B7: It is a novel depth-enhanced hole filling approach for DIBR view interpolation [68]. Instead of inpainting the warped images directly, it focuses on the use of the

occluded information to identify the relevant background pixels around the holes.

Compared with the IRCCyN/IVC database, the synthesized views in IETR database do not include geometric distortions (black holes), since the IETR database adopts the DIBR algorithms with better performance.

### B. Competing IQA Metrics and Evaluation Methodology

In order to check the superiority and validity of the proposed metric, we compare our proposed model with the recently developed IQA metrics, which can be divided into two categories. The first category of IQA algorithms were designed for natural scene images, including the following:

- PSNR: It computes the differences of pixel level.
- SSIM: It combines the deviations of luminance, contrast and structural similarity between a distorted image and its associated reference image [4].
- IW-SSIM: This method is based on the statistical information theory and the NSS module, and uses the information weighting strategy to weight the SSIM. [69].
- FSIM: It calculates the similarity of phase congruency and gradient magnitude between an reference image and its associated distorted image [70].
- ADD-SSIM: By analyzing the distortion distribution affected by image content and distortion, an effective pooling scheme is proposed to improve the performance of SSIM [71].
- PSIM: It combines the gradient magnitude similarities at two scales, the color information similarity, and a reliable perceptual-based pooling [72].
- NIQE: Based on a simple and successful space domain NSS model, a statistical features set of “quality aware” is constructed. The quality of distorted images is expressed as a simple distance measure between model statistics and those distorted images.[18].
- IL-NIQE: The algorithm takes more features into account on the basis of NIQE [19].
- ARISM: It analysis the parameters of classical AR to evaluate image quality [14].
- BIQME: It considers five influencing factors, including image contrast, sharpness, brightness, colorfulness, and naturalness, which blindly predict visual quality [73].

In addition, we further compare our proposed blind quality measurement method with the existing reference-based and reference-free IQA methods intended for DIBR-synthesized images. These methods in the second category include the followings:

- MW-PSNR: It deploys morphological wavelet decomposition to evaluate the quality of DIBR-synthesized views [22].
- MP-PSNR: It is similar to MW-PSNR, but the morphologic wavelet decomposition is replaced by the morphological pyramid decomposition [23].
- MP-PSNR-reduce: It modifies MP-PSNR to obtain better performance and faster computation [24].
- LOGS: It measures the local geometric distortions in disoccluded regions and global sharpness [26].

- NIQSV+: It can evaluate the quality of synthesized views by measuring the typical synthesis distortions: blurry regions, black holes, and stretching [31].
- APT: It is a new NSS module based on AR-based local image description [32].
- CLGM: It combines local and global measures to assess three DIBR-related distortions including disoccluded region, stretching and global sharpness [33].
- OUT: It uses the median filtering to highlight both geometric and structural distortions and further fused these highlighted distortions to predict the quality of DIBR-synthesized views [34].

In order to evaluate the performance of IQA metrics, we have adopted three commonly employed criteria, including Spearman Rank order Correlation Coefficient (SRCC), Pearson Linear Correlation Coefficient (PLCC) and Root Mean Square Error (RMSE). SRCC is employed to measure the prediction monotonicity. PLCC and RMSE are used to measure the prediction accuracy. The better IQA metrics should achieve higher values of SRCC and PLCC, but obtain a lower value of RMSE. They are calculated following a five-parameter nonlinear mapping:

$$f(x) = \tau_1 \left( \frac{1}{2} - \frac{1}{1 + e^{\tau_2(x - \tau_3)}} \right) + \tau_4 x + \tau_5, \quad (14)$$

where  $\tau_i, i = 1, 2, 3, 4, 5$  are the parameters to be fitted;  $x$  and  $f(x)$  represent the objective scores predicted by IQA metrics and its associated subjective scores.

### C. Performance Comparison

Table I shows the comparison between the proposed quality metric with the existing IQA metrics including both the first and second categories. The full-reference LOGS method obtains the best overall performance as compared with other competing IQA algorithms on the IRCCyN/IVC and IETR datasets. The performances of our proposed algorithm on IRCCyN/IVC and IETR datasets are superior to the competing reference-free IQA models designed for natural scene images and DIBR-synthesized views. Moreover, the performance of our proposed algorithm on the IRCCyN/IVC database is second only to that of the LOGS method. We can find two important conclusions from Table I. First, the existing IQA metrics in the first category, designed for natural scene images, have difficulty in evaluating DIBR-synthesized images. In the IRCCyN/IVC database, the BIQME method in the first type of IQA metrics leads to the optimal results and obtains 0.6770, 0.7271 and 0.4571 of SRCC, PLCC and RMSE, respectively, while the value of SRCC still does not exceed 0.7. In the IETR database, the IW-SSIM method contained in the first type of IQA models achieves the 0.5950, 0.6282 and 0.1933 of SRCC, PLCC and RMSE, respectively, which are obviously inferior to the SRCC, PLCC and RMSE values of the LOGS method. Experiments confirm our view that the first type of IQA algorithms cannot effectively predict the quality of DIBR-synthesized views. Second, since the synthesized views in IETR database do not include obvious geometric distortions, the performances of the proposed algorithm and competing

TABLE I  
PERFORMANCE COMPARISON OF STATE-OF-THE-ART IQA MODELS ON THE IRCCyN/IVC AND IETR DIBR DATABASES

Metric	Type	IRCCyN/IVC Database [1]			IETR Database [2]		
		SRCC	PLCC	RMSE	SRCC	PLCC	RMSE
PSNR	Reference-based	0.3095	0.3976	0.6109	0.5356	0.6012	0.1985
SSIM [4]	Reference-based	0.4368	0.4850	0.5823	0.2395	0.4016	0.2275
IW-SSIM [69]	Reference-based	0.4053	<b>0.5831</b>	<b>0.5409</b>	<b>0.5950</b>	<b>0.6280</b>	<b>0.1933</b>
FSIM [70]	Reference-based	0.4148	0.5828	0.5411	0.5250	0.5785	0.2022
ADD-SSIM [71]	Reference-based	<b>0.4672</b>	0.5512	0.5556	0.5248	0.4853	0.2168
PSIM [72]	Reference-based	0.4576	0.5315	0.5640	0.5030	0.5371	0.2091
NIQE [18]	Reference-free	0.3739	0.4374	0.5987	0.1360	0.2244	0.2421
IL-NIQE [19]	Reference-free	0.5348	0.4998	0.5767	0.1867	0.1728	0.2442
ARISM [14]	Reference-free	0.3728	0.3994	0.6104	<b>0.2035</b>	<b>0.1984</b>	<b>0.2430</b>
BIQME [73]	Reference-free	<b>0.6770</b>	<b>0.7271</b>	<b>0.4571</b>	0.0701	0.1168	0.2462
MW-PSNR [22]	Reference-based	0.5757	0.5622	0.5506	0.4875	0.5389	0.2088
MP-PSNR [23]	Reference-based	0.6227	0.6174	0.5238	0.5809	0.6190	0.1947
MP-PSNR-reduce[24]	Reference-based	0.6634	0.6772	0.4899	0.5870	0.6160	0.1953
LOGS [26]	Reference-based	<b>0.7812</b>	<b>0.8256</b>	<b>0.3601</b>	<b>0.6679</b>	<b>0.6638</b>	<b>0.1854</b>
NIQSV+ [31]	Reference-free	0.6668	0.7114	0.4679	0.2190	0.2095	0.2429
APT [32]	Reference-free	0.7157	0.7307	0.4546	0.4141	0.4225	0.2252
CLGM [33]	Reference-free	0.6528	0.6750	0.4620	0.0860	0.1146	0.2463
OUT [34]	Reference-free	0.7036	0.7678	0.4266	0.2378	0.2409	0.2406
Proposed	Reference-free	<b>0.7867</b>	<b>0.7995</b>	<b>0.4000</b>	<b>0.4338</b>	<b>0.4254</b>	<b>0.2244</b>

IQA metrics on IETR database is not well. The performance of the LOGS method on IRCCyN/IVC and IETR databases are superior to the reference-free IQA methods designed for DIBR-synthesized views. Since the reference images are generally not accessible in most real applications, it is of great significance to study the reference-free IQA models for DIBR-synthesized views. In reference-free IQA metrics devised for DIBR-synthesized views, the proposed method has achieved the best performance as compared with competing IQA metrics on IRCCyN/IVC and IETR databases. For a more intuitive comparison, Fig. 4 provides scatter diagrams of subjective DMOS values versus objective quality scores estimated by different IQA metrics, which include: MP-PSNR, MP-PSNR-reduce, LOGS, NIQSV+, APT, CLGM, OUT and our proposed algorithm designed for DIBR-synthesized images. It is obvious that the sample points of our proposed algorithm present higher convergence and linearity as compared with other competing reference-free IQA metrics devised for synthesized views. It means objective scores predicted by the proposed metric maintain more consistency in line with subjective ratings.

In addition, Table II further shows the performance of  $Q_1$ ,  $Q_2$ ,  $Q_1/Q_3$  and  $Q_2/Q_3$  on the IRCCyN/IVC database. As shown in Table II, since the performances of  $Q_1/Q_3$  and  $Q_2/Q_3$  on IRCCyN/IVC dataset have been greatly improved compared to  $Q_1$  and  $Q_3$ , the image complexity  $Q_3$  can effectively weaken the negative impact of image content diversity on the proposed reference-free IQA model. Compared to

TABLE II  
THE PERFORMANCE OF  $Q_1$ ,  $Q_2$ ,  $Q_1/Q_3$  AND  $Q_2/Q_3$  ON THE IRCCyN/IVC DATABASE

Metric	$Q_1$	$Q_2$	$Q_1/Q_3$	$Q_2/Q_3$
PLCC	0.6903	0.6512	0.7332	0.7217
SRCC	0.6616	0.4573	0.7551	0.6156
RMSE	0.4817	0.5053	0.4528	0.4609

the performance of the overall model on the IRCCyN/IVC database shown in Table I, the  $Q_1$  and  $Q_2$  models proposed in this research can complement each other to evaluate DIBR-synthesized views.

In order to further check the effectiveness of our proposed method on different DIBR algorithms, the IRCCyN/IVC [1] and IETR databases [2] are divided into 7 parts, i.e. A1-A7 and B1-B7. Table III shows the experimental results evaluated by the SRCC for each part. Obviously, our proposed metric achieves excellent performance for the algorithms A5-A7 and B7, while fails to evaluate the algorithms A1-A4 and B1-B6. This phenomenon is probably due to the fact that the views synthesized by the algorithms A1-A4 and B1-B6 have almost little geometric and blur distortions.

#### D. Comparison on Statistical Significance

In addition to the SRCC, PLCC and RMSE employed for numerical comparisons, more attention was paid to the statistical significance comparison recently.  $F_{test}$  is a typical

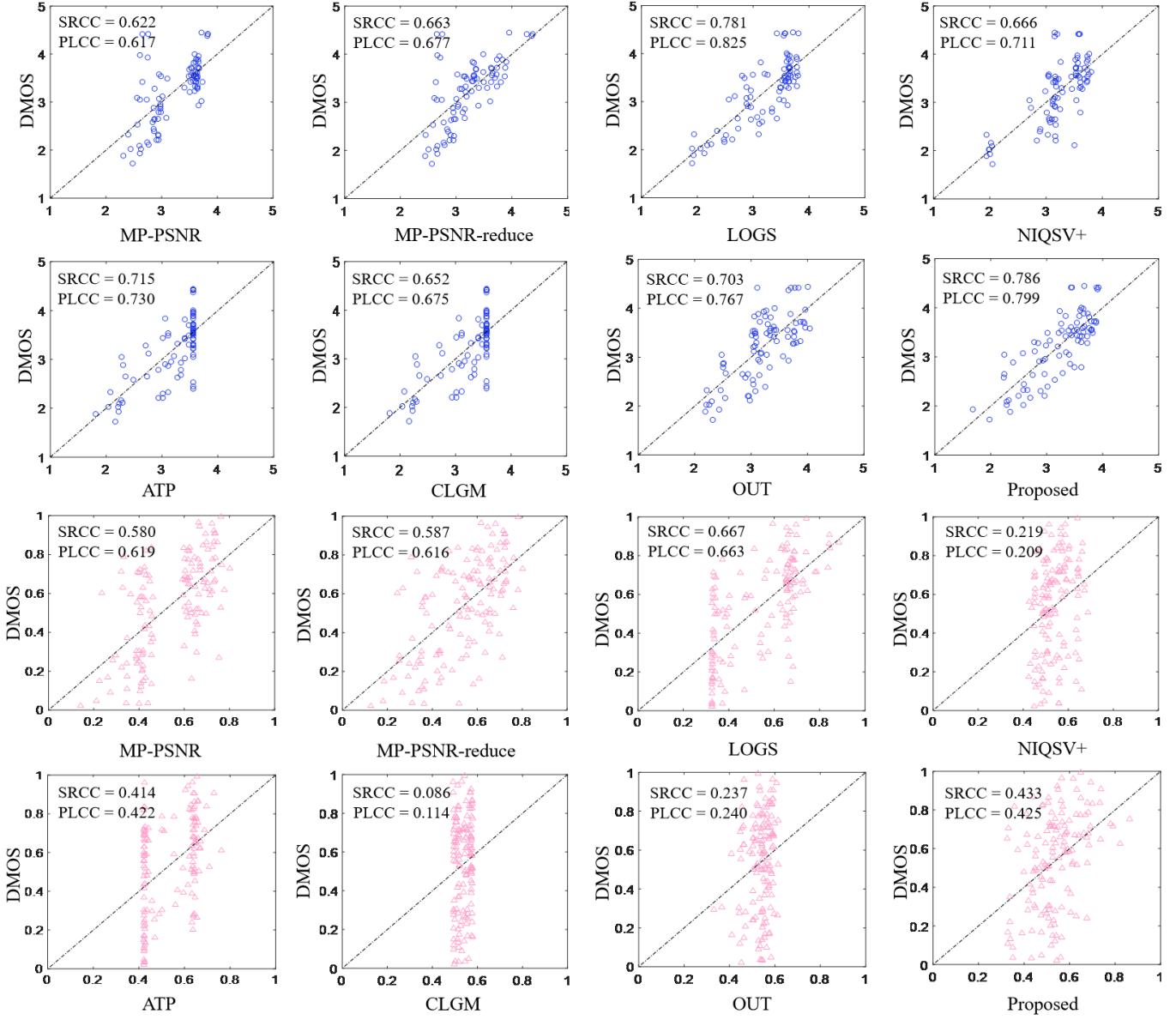


Fig. 4. Scatter diagrams of DMOS versus reference-based (MP-PSNR, MP-PSNR-reduce and LOGS) and reference-free(NIQSV+, APT, CLGM, OUT and proposed algorithm) models designed for DIBR-synthesized views on the IRCCyN/IVC (first two lines) and IETR (last two lines) databases. Since we pay more attention on comparisons with reference-free methods, only partial reference-based methods are involved here. The black diagonal dash line means the perfect prediction.

TABLE III

RANK AND SRCC OF DIBR METHODS A1-A7 AND B1-B7 CONTAINED IN IRCCyN/IVC and IETR Databases

Method	A1	A2	A3	A4	A5	A6	A7
Rank	6	5	4	7	2	3	1
SRCC	0.50	0.68	0.70	0.49	0.86	0.85	0.94
Method	B1	B2	B3	B4	B5	B6	B7
Rank	4	6	2	3	5	7	1
SRCC	0.51	0.33	0.66	0.60	0.41	0.03	0.83

statistical test that follows the variance-based hypothesis and reveals additional information about the relative performance of different visual quality assessment models [74].

$F_{test}$  is calculated based on the residual differences of subjective quality ratings and the associated objective quality predictions. Specifically, the  $F_{test}$  between the proposed algorithm and a competing algorithm is defined as follow

$$F_{test} = \sigma_c^2 / \sigma_p^2, \quad (15)$$

where  $\sigma_c$  and  $\sigma_p$  represent the RMSE values of a competing method and the proposed method tested on the DIBR-synthesized image database, respectively. We obtain the statistical significance between all competing methods and the proposed metric based on comparing  $F_{test}$  with a threshold  $F_{critical}$ , which is determined by the confidence level and the number of images in the database. In particular, if  $F_{test}$  is bigger (smaller) than  $F_{critical}$  ( $1/F_{critical}$ ), the performance of the proposed algorithm is statistically better (worse) than

	M1	M2	M3	M4	M5	M6	M7	M8	M9
M1	-	+1	+1	+1	0	0	0	0	0
M2	-1	-	0	0	-1	0	-1	0	-1
M3	-1	0	-	0	-1	0	0	0	-1
M4	-1	0	0	-	-1	0	0	0	0
M5	0	+1	+1	+1	-	+1	+1	+1	-
M6	0	0	0	0	-1	-	0	0	0
M7	0	+1	0	0	-1	0	-	0	0
M8	0	0	0	0	-1	0	0	-	0
M9	0	+1	+1	0	0	0	0	0	-

Fig. 5. Statistical significance comparison between the proposed algorithm and competing IQA methods designed for DIBR-synthesized images on the IRCCyN/IVC database. M1: Proposed algorithm; M2: MW-PSNR; M3: MP-PSNR; M4: MP-PSNR-reduce; M5: LOGS; M6: NIQSV+; M7: APT; M8: CLGM; M9: OUT.

the competing models. When  $1/F_{critical} < F_{test} < F_{critical}$ , the two quality methods are statistically competitive. The results of the  $F_{test}$  statistics on the IRCCyN/IVC and IETR databases are shown in Fig. 5 and Fig. 6. A value of ‘0’ indicates that the two quality models are statistically equally competitive, while ‘+1’ and ‘-1’ respectively indicate that the method in the row are statistically superior and inferior to that in the column. To make the results clearer, we further use three different colors to mark these results. From the Fig. 5, we can find that the performance of the proposed IQA method on IRCCyN/IVC database is significantly superior to the reference-based IQA algorithms designed for synthesized views, and has no significant difference compared to the tested reference-free DIBR-synthesized image quality assessment methods. Fig. 6 shows the performance of the proposed IQA algorithm on the IETR database has no significant difference compared to the tested reference-based and reference-free IQA methods devised for DIBR-synthesized views.

#### E. Performance Dependency of Used Parameters

A stable IQA model is more desirable, that is, the performance of a fine IQA model should not vary significantly with the minor change of the used parameters. In this subsection, we discuss the impact of the parameter  $\alpha$  used in equation (12) on the performance of our proposed metric. The  $\alpha$  is used to adjust the contributions of  $Q_1$  and  $Q_2$ . Since geometric distortion seriously destroys the visual quality of DIBR-synthesized views, we surmise model  $Q_1$  should account for more proportion than model  $Q_2$ . We further examine the dependency of the proposed model on the parameter  $\alpha$  and offer the results in Fig. 7, in which two important conclusions can be observed below.

- With the increase of the parameter  $\alpha$ , the performance of our proposed algorithm does not change greatly.

	M1	M2	M3	M4	M5	M6	M7	M8	M9
M1	-	0	-1	0	-1	0	0	0	0
M2	0	-	0	0	0	+1	0	+1	+1
M3	+1	0	-	0	0	+1	+1	+1	+1
M4	0	0	0	-	0	+1	+1	+1	+1
M5	+1	0	0	0	-	+1	+1	+1	+1
M6	0	-1	-1	-1	-1	-	0	0	0
M7	0	0	-1	-1	-1	0	-	0	0
M8	0	-1	-1	-1	-1	0	0	-	0
M9	0	-1	-1	-1	-1	0	0	0	-

Fig. 6. Statistical significance comparison between the proposed algorithm and competing IQA methods designed for DIBR-synthesized images on the IETR database. M1: Proposed algorithm; M2: MW-PSNR; M3: MP-PSNR; M4: MP-PSNR-reduce; M5: LOGS; M6: NIQSV+; M7: APT; M8: CLGM; M9: OUT.

The worst performance of our proposed method has obtained the 0.6239, 0.6493 and 0.5064 of SRCC, PLCC and RMSE, respectively, which are still superior over the most state-of-the-art IQA metrics tested.

- 2) The results show that the performance of the algorithm is better when the parameter  $\alpha$  is less than 0.5. It also proves that the evaluation of geometric distortions in DIBR-synthesized views should be dominant in our algorithm. Considering the values of SRCC, PLCC and RMSE, we set the parameter  $\alpha = 0.15$  to optimize the performance of the proposed algorithm.

#### F. Implementation Time

DIBR technologies are employed to synthesize FVV images in the “blind” environment. Therefore, a real-time and reliable blind assessment model is particularly favorite. The computational efficiency becomes another key factor, which should be taken consideration into the process of designing IQA metrics. In this subsection, we conduct an experiment on the IRCCyN/IVC database to investigate the implementation time of our proposed metric and its competing reference-free IQA metrics. All the IQA metrics are performed on the same test environment (MATLAB R2017a software executed on a 2.3 GHz processor with 8GB RAM, Windows 10 Pro 64-bit desktop). Table IV exhibits the processing time of each image (in term of seconds) of nine reference-free IQA algorithms. It is apparent that OUT is the most time-saving method, while its performance is the worse than the proposed metric. APT costs more than one minute, which far exceeds the running time of other reference-free IQA algorithms designed for synthesized views. The running time of the proposed method is slower than the NIQSV+, CLGM and OUT models, but its performance is optimal in the reference-free IQA metrics devised for natural scene and synthesized images.

TABLE IV  
THE IMPLEMENTATION TIME OF REFERENCE-FREE IQA MODELS (IN SECONDS)

Metric	NIQE	IL-NIQE	ARISM	BIQME	NIQSV+	APT	CLGM	OUT	Proposed
Time(s)	0.50	3.97	30.84	0.86	0.23	95.31	0.31	0.11	1.42

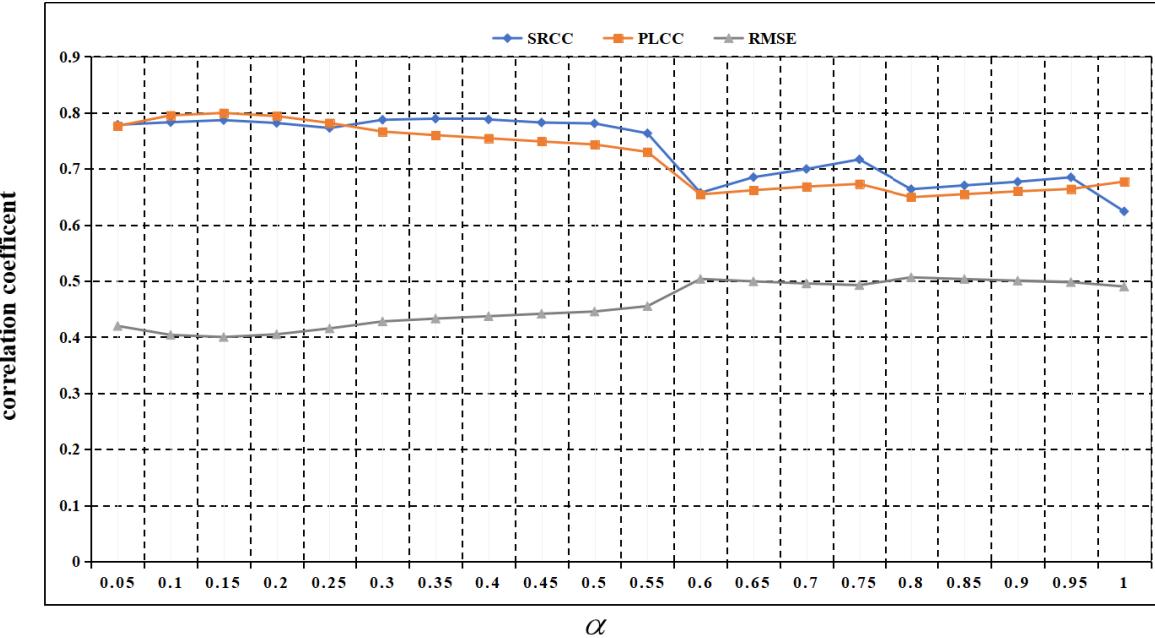


Fig. 7. Performance dependency of the proposed metric with the changing parameter  $\alpha$ .

#### IV. CONCLUSION

The DIBR-synthesized images are the basis of free-view point videos, virtual reality and augmented reality, etc. In the literature, some methods were proposed to produce the DIBR-synthesized views but highly coherent and fast blind IQA metrics designed for DIBR-synthesized images are missing. In this paper, we have proposed a novel blind IQA metric, which contains some spotlights as compared with existing IQA algorithms. First, we measure the geometric distortion and global sharpness in the discrete wavelet transform domain. Then, a hybrid filter combining AR and BL filters is employed to compute the image complexity, which is used to normalize the geometric distortion and global sharpness to achieve the final quality score. Second, the performance of the proposed method is superior to the competing reference-free IQA metrics. Third, our proposed method is sensitive to the parameter  $\alpha$ , but its worst performance also reaches the mainstream level of the competing IQA algorithms. Fourth, the computational requirements of our proposed algorithm has reached the mainstream level. Experiments conducted on IRCCyN/IVC and IETR databases show the superiority of our proposed algorithm as compared with competing state-of-the-art reference-free IQA metrics, which were devised for natural scene and DIBR-synthesized images. Due to the fact that geometric distortion scarcely appears in high-quality DIBR-synthesized images, our proposed model mistakes these images and leads to the poor correlation performance. In the

future work, we will consider how to effectively evaluate high-quality DIBR-synthesized views and reduce the computational requirements of IQA model, as well as propose quality models for multiple synthesis of close viewpoints.

#### REFERENCES

- [1] E. Bosc *et al.*, “Towards a new quality metric for 3-D synthesized view assessment,” *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 7, pp. 1332–1343, Nov. 2011.
- [2] S. Tian, L. Zhang, L. Morin, and O. Déforges, “A benchmark of DIBR synthesized view quality assessment metrics on a new database for immersive media applications,” *IEEE Trans. Multimedia*, vol. 21, no. 5, pp. 1235–1247, May 2019.
- [3] B. Wang, Z. Wang, Y. Liao, and X. Lin, “HVS-based structural similarity for image quality assessment,” in *Proc. 9th Int. Conf. Signal Process.*, 2008, pp. 1194–1197.
- [4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [5] K. Gu, G. Zhai, X. Yang, and W. Zhang, “Using free energy principle for blind image quality assessment,” *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 50–63, Jan. 2015.
- [6] Z. Shao and J. Cai, “Remote sensing image fusion with deep convolutional neural network,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1656–1669, May 2018.
- [7] L. Zhou, Z. Wang, Y. Luo, and Z. Xiong, “Separability and compactness network for image recognition and superresolution,” *IEEE Trans. Neural Netw. Learn. Syst.*, to be published.
- [8] Z. Wang *et al.*, “Multi-memory convolutional neural network for video super-resolution,” *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 2530–2544, May 2019.
- [9] K. Jiang, Z. Wang, P. Yi, G. Wang, T. Lu, and J. Jiang, “Edge-enhanced GAN for remote sensing image superresolution,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5799–5812, Aug. 2019.

- [10] P. Yi, Z. Wang, K. Jiang, Z. Shao, and J. Ma, "Multi-temporal ultra dense memory network for video super-resolution," *IEEE Trans. Circuits Syst. Video Technol.*, to be published.
- [11] S. Zhenfeng, L. Jun, and C. Qimin, "Fusion of infrared and visible images based on focus measure operators in the curvelet domain," *Appl. Opt.*, vol. 51, no. 12, pp. 1910–1921, Apr. 2012.
- [12] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.
- [13] C. T. Vu, T. D. Phan, and D. M. Chandler, " $S_3$ : A spectral and spatial measure of local perceived sharpness in natural images," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 934–945, Mar. 2012.
- [14] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3218–3231, Oct. 2015.
- [15] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352, Aug. 2012.
- [16] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.
- [17] P. Ye, J. Kumar, L. Kang, and D. Doermann, "Unsupervised feature learning framework for no-reference image quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 1098–1105.
- [18] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.
- [19] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015.
- [20] Y. Zhou, L. Li, J. Wu, K. Gu, W. Dong, and G. Shi, "Blind quality index for multiply distorted images using biorider structure degradation and nonlocal statistics," *IEEE Trans. Multimedia*, vol. 20, no. 11, pp. 3019–3032, Nov. 2018.
- [21] F. Battisti, E. Bosc, M. Carli, P. Le Callet, and S. Perugia, "Objective image quality assessment of 3D synthesized views," *Signal Process., Image Commun.*, vol. 30, pp. 78–88, Jan. 2015.
- [22] D. Sandić-Stanković, D. Kukolj, and P. Le Callet, "DIBR synthesized image quality assessment based on morphological wavelets," in *Proc. 7th Int. Workshop Qual. Multimedia Exper.*, Jan. 2015, pp. 1–6.
- [23] D. Sandić-Stanković, D. Kukolj, and P. Le Callet, "DIBR-synthesized image quality assessment based on morphological multi-scale approach," in *Proc. True Vis.-Capture, Transmiss. Display 3D Video*, vol. 1, Oct. 2015, pp. 1–4.
- [24] D. Sandić-Stanković, D. Kukolj, and P. Le Callet, "Multi-scale synthesized view assessment based on morphological pyramids," *J. Elect. Eng.*, vol. 67, no. 1, pp. 1–9, Jan. 2016.
- [25] V. Jakhetiya, K. Gu, W. Lin, Q. Li, and S. P. Jaiswal, "A prediction backed model for quality assessment of screen content and 3-D synthesized images," *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 652–660, Feb. 2018.
- [26] L. D. Li, Y. Zhou, K. Gu, W. S. Lin, and S. Q. Wang, "Quality assessment of DIBR-synthesized images by measuring local geometric distortions and global sharpness," *IEEE Trans. Multimedia*, vol. 20, no. 4, pp. 914–926, Apr. 2018.
- [27] M. S. Farid, M. Lucenteforte, and M. Grangetto, "Perceptual quality assessment of 3D synthesized images," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2017, pp. 505–510.
- [28] S. Ling and P. Le Callet, "How to learn the effect of non-uniform distortion on perceived visual quality? Case study using convolutional sparse coding for quality assessment of synthesized views," in *Proc. 25th IEEE Int. Conf. Image Process.*, Jul. 2018, pp. 286–290.
- [29] S. Ling, J. Gutiérrez, K. Gu, and P. Le Callet, "Prediction of the influence of navigation scan-path on perceived quality of free-viewpoint videos," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 1, pp. 204–216, Mar. 2019.
- [30] S. Tian, L. Zhang, L. Morin, and O. Déforges, "NIQSV: A no reference image quality assessment metric for 3D synthesized views," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2017, pp. 1248–1252.
- [31] S. Tian, L. Zhang, L. Morin, and O. Déforges, "NIQSV+: A no-reference synthesized view quality assessment metric," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 1652–1664, Apr. 2018.
- [32] K. Gu, V. Jakhetiya, J.-F. Qiao, X. Li, W. Lin, and D. Thalmann, "Model-based referenceless quality metric of 3D synthesized images using local image description," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 394–405, Jan. 2018.
- [33] G. Yue, C. Hou, K. Gu, T. Zhou, and G. Zhai, "Combining local and global measures for DIBR-synthesized image quality evaluation," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 2075–2088, Apr. 2019.
- [34] V. Jakhetiya, K. Gu, T. Singhal, S. C. Guntuku, Z. Xia, and W. Lin, "A highly efficient blind image quality assessment metric of 3-D synthesized images using outlier detection," *IEEE Trans. Ind. Informat.*, vol. 15, no. 7, pp. 4120–4128, Jul. 2019.
- [35] Y. Zhou, L. Li, S. Wang, J. Wu, Y. Fang, and X. Gao, "No-reference quality assessment for view synthesis using DoG-based edge statistics and texture naturalness," *IEEE Trans. Image Process.*, vol. 28, no. 9, pp. 4566–4579, Sep. 2019.
- [36] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 45, no. 5, pp. 485–560, 1992.
- [37] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
- [38] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
- [39] P. Merkle, A. Smolic, K. Muller, and T. Wiegand, "Multi-view video plus depth representation and coding," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, Sep. 2007, pp. I-201–I-204.
- [40] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 717–728, Apr. 2009.
- [41] N. D. Narvekar and L. J. Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2678–2683, Sep. 2011.
- [42] R. Hassen, Z. Wang, and M. M. A. Salama, "Image sharpness assessment based on local phase coherence," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2798–2810, Jul. 2013.
- [43] L. Li, W. Lin, X. Wang, G. Yang, K. Bahrami, and A. C. Kot, "No-reference image blur assessment based on discrete orthogonal moments," *IEEE Trans. Cybern.*, vol. 46, no. 1, pp. 39–50, Jan. 2016.
- [44] P. V. Vu and D. M. Chandler, "A fast wavelet-based algorithm for global and local image sharpness estimation," *IEEE Signal Process. Lett.*, vol. 19, no. 7, pp. 423–426, Jul. 2012.
- [45] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, "No-reference quality metric of contrast-distorted images based on information maximization," *IEEE Trans. Cybern.*, vol. 47, no. 12, pp. 4559–4565, Dec. 2017.
- [46] K. Gu, J. Zhou, J.-F. Qiao, G. Zhai, W. Lin, and A. C. Bovik, "No-reference quality assessment of screen content pictures," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 4005–4018, Aug. 2017.
- [47] K. Gu, J. Qiao, X. Min, G. Yue, W. Lin, and D. Thalmann, "Evaluating quality of screen content images via structural variation analysis," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 10, pp. 2689–2701, Oct. 2018.
- [48] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Visual saliency detection with free energy theory," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1552–1555, Oct. 2015.
- [49] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE 6th Int. Conf. Comput. Vis.*, Jan. 1998, pp. 836–846.
- [50] K. Gu, J. Qiao, and X. Li, "Highly efficient picture-based prediction of PM 2.5 concentration," *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 3176–3184, Apr. 2019.
- [51] A. Mittal, M. A. Saad, and A. C. Bovik, "A completely blind video integrity oracle," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 289–300, Jan. 2016.
- [52] K. Gu, J. Qiao, S. Lee, H. Liu, W. Lin, and P. Le Callet, "Multiscale natural scene statistical analysis for no-reference quality evaluation of DIBR-synthesized views," *IEEE Trans. Broadcast.*, to be published.
- [53] C. Fehn, "Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV," *Proc. SPIE*, vol. 5291, no. 2, pp. 93–104, May 2004.
- [54] A. Telea, "An image inpainting technique based on the fast marching method," *J. Graph. Tools*, vol. 9, no. 1, pp. 23–34, 2004.
- [55] Y. Mori, N. Fukushima, T. Yendo, T. Fujii, and M. Tanimoto, "View generation with 3D warping using depth information for FTV," *Signal Process., Image Commun.*, vol. 24, nos. 1–2, pp. 65–72, Jan. 2009.

- [56] K. Müller, A. Smolic, K. Dix, P. Merkle, P. Kauff, and T. Wiegand, "View synthesis for advanced 3D video systems," *EURASIP J. Image Video Process.*, vol. 2008, no. 1, 2009, Art. no. 438148.
- [57] P. Ndjiki-Nya *et al.*, "Depth image based rendering with advanced texture synthesis," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2010, pp. 424–429.
- [58] M. Köppel *et al.*, "Temporally consistent handling of disocclusions with texture synthesis for depth-image-based rendering," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 1809–1812.
- [59] *Subjective Video Quality Assessment Methods for Multimedia Applications*, document ITU-T study group 12, ITU-T p.910, 1997.
- [60] G. Cermak, L. Thorpe, and M. Pinson, "Test plan for evaluation of video quality models for use with high definition TV content," in *Video Quality Experts Group*, 2009.
- [61] J. L. Blin, "New quality evaluation method suited to multimedia context: SAMVIQ," in *Proc. Int. Workshop Video Process. Quality Metrics*, 2006, pp. 1–4.
- [62] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1200–1212, Sep. 2004.
- [63] V. Jantet, C. Guillemot, and L. Morin, "Object-based layered depth images for improved virtual view synthesis in rate-constrained context," in *Proc. IEEE Conf. Image Process.*, Sep. 2011, pp. 125–128.
- [64] I. Ahi and C. Kim, "A novel depth-based virtual view synthesis method for free viewpoint video," *IEEE Trans. Broadcast.*, vol. 59, no. 4, pp. 614–626, Dec. 2013.
- [65] G. Luo, Y. Zhu, Z. Li, and L. Zhang, "A hole filling approach based on background reconstruction for view synthesis in 3D video," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1781–1789.
- [66] M. Soll and G. AlRegib, "Hierarchical hole-filling for depth-based view synthesis in FTV and 3D video," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 5, pp. 495–504, Sep. 2012.
- [67] M. Tanimoto, T. Fujii, K. Suzuki, N. Fukushima, and Y. Mori, *Reference Softwares for Depth Estimation and View Synthesis*, document M15377, ISO/IECJTC1/SC29/WG11 MPEG 20081, 2008.
- [68] C. Zhu and S. Li, "Depth image based view synthesis: New insights and perspectives on hole generation and filling," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 82–93, Mar. 2016.
- [69] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1185–1198, May 2011.
- [70] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
- [71] K. Gu, S. Wang, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Analysis of distortion distribution for pooling in image quality prediction," *IEEE Trans. Broadcast.*, vol. 62, no. 2, pp. 446–456, Jun. 2016.
- [72] K. Gu, L. Li, H. Lu, X. Min, and W. Lin, "A fast reliable image quality predictor by fusing micro- and macro-structures," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 3903–3912, May 2017.
- [73] K. Gu, D. Tao, J.-F. Qiao, and W. Lin, "Learning a no-reference quality assessment model of enhanced images with big data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 4, pp. 1301–1313, Apr. 2018.
- [74] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.



**Guangcheng Wang** received the B.S. degree from the School of Electrical and Optoelectronic Engineering, Changzhou Institute of Technology, Changzhou, China, in 2015, and the M.S. degree from the School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, China, in 2018. He is currently pursuing the Ph.D. degree with the National Engineering Research Center for Multimedia Software, School of Computer Science, Wuhan University, Wuhan, China. His research interests include image processing, quality assessment, and face-anti spoofing.



**Zhongyuan Wang** received the Ph.D. degree in communication and information system from Wuhan University, Wuhan, China, in 2008. He is currently a Professor with the School of Computer Science, Wuhan University. He is currently directing two projects funded by the National Natural Science Foundation Program of China. His research interests include video compression, image processing, and multimedia communications.



**Ke Gu** (M'19) received the B.S. and Ph.D. degrees in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2009 and 2015, respectively. He is currently a Professor with the Beijing University of Technology, Beijing, China. His research interests include environmental perception, image processing, quality assessment, and machine learning. He was a recipient of the Best Paper Award from the IEEE TRANSACTIONS ON MULTIMEDIA (T-MM), the Best Student Paper Award at the IEEE International Conference on Multimedia and Expo (ICME) in 2016, and the Excellent Ph.D. Thesis Award from the Chinese Institute of Electronics in 2016. He was the Leading Special Session Organizer in the VCIP 2016 and the ICIP 2017, and serves as a Guest Editor of *Digital Signal Processing* (DSP). He is currently an Associate Editor of IEEE ACCESS and *IET Image Processing* (IET-IPR), and an Area Editor of *Signal Processing Image Communication* (SPIC). He is a reviewer of 20 top SCI journals.



**Leida Li** received the B.S. and Ph.D. degrees from Xidian University, Xian, China, in 2004 and 2009, respectively. In 2008, he was a Research Assistant with the Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Taiwan. From 2014 to 2015, he was a Visiting Research Fellow with the Rapid-Rich Object Search (ROSE) Laboratory, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, where he was a Senior Research Fellow from 2016 to 2017. He is currently a Full Professor with the School of Artificial Intelligence, Xidian University. His current research interests include multimedia quality assessment, affective computing, information hiding, and image forensics. He is an Associate Editor of the *Journal of Visual Communication and Image Representation* and the *EURASIP Journal on Image and Video Processing*.



a Full Professor with the School of Artificial Intelligence, Xidian University. His current research interests include multimedia quality assessment, affective computing, information hiding, and image forensics. He is an Associate Editor of the *Journal of Visual Communication and Image Representation* and the *EURASIP Journal on Image and Video Processing*.

**Zhifang Xia** received the B.S. degree in measuring and control instrument from Anhui University, Hefei, China, in 2008, and the master's degree in control science and engineering from Tsinghua University, Beijing, China, in 2012. She is currently pursuing the Ph.D. degree with the Beijing University of Technology, Beijing. She is currently an Engineer and a registered Consultant (investment) with State Information Center, Beijing. Her research interests include image processing, quality assessment, machine learning, and e-government. She was a recipient of the Second Prize of the National Excellent Engineering Consultation Award in 2016.



**Lifang Wu** received the B.E. and M.E. degrees and the Ph.D. degree in pattern recognition and intelligent system from the Beijing University of Technology (BJUT), Beijing, China, in 1991, 1994, and 2003, respectively. She is currently a faculty of the School of Electronic Information and Control Engineering, BJUT, where she currently serves as a Professor. She has published over 50 referred technical articles in international journals and conferences of image and video processing, and pattern recognition. Her current research interests include image and video analysis and understanding, face detection and recognition, and face encryption. She is a Senior Member of the Chinese Institute of Electronics.