BLIND QUALITY ASSESSMENT FOR 3D-SYNTHESIZED IMAGES BY MEASURING GEOMETRIC DISTORTIONS AND IMAGE COMPLEXITY

Guangcheng Wang[†], *Zhongyuan Wang*[†], *Ke Gu*[‡], *and Zhifang Xia*[⋆]

[†]National Engineering Research Center for Multimedia Software, School of Computer Science, Wuhan University, China
[‡]Faculty of Information Technology, Beijing University of Technology, China
[⋆]The State Information Center of P.R.China, Beijing, China

ABSTRACT

Free viewpoint video (FVV), owing to its comprehensive applications in immersive entertainment, remote surveillance and distanced education, has received extensive attention and been regarded as a new important direction of video technology development. Depth image-based rendering (DIBR) technologies are employed to synthesize FVV images in the "blind" environment. Therefore, a real-time reliable blind quality assessment metric is urgently required. However, existing stste-of-art quality assessment methods are limited to estimate geometric distortions generated by DIBR. In this research, a novel blind quality metric, measuring Geometric Distortions and Image Complexity (GDIC), is proposed for DIBR-synthesized images. Firstly, a DIBR-synthesized image is decomposed into wavelet subbands by using discrete wavelet transform. Then, we adopt canny operator to capture the edge of wavelet subbands and compute the edge similarity between low-frequency subband and high-frequency subbands. The edge similarity is used to quantify geometric distortions in DIBR-synthesized images. Secondly, a hybrid filter combining the autoregressive and bilateral filter is adopted to compute image complexity. Finally, the overall quality score is calculated by normalizing geometric distortions via image complexity. Experiments show that our proposed GDIC is superior to prevailing image quality assessment metrics, which were intended for natural and DIBR-synthesized images.

Index Terms— Depth image-based rendering, image quality assessment, blind

1. INTRODUCTION

With the rapid development of 3D-related technologies, many new challenges have emerged. Free viewpoint video (FVV), owing to its comprehensive applications in immersive entertainment, remote surveillance and distanced education, has received extensive attention and been regarded as a new important direction of video technology development. FVV promotes the rapid development of Depth Image-Based Rendering (DIBR) technology, which can produce new viewpoints from multiple views [1]. Unfortunately, these new viewpoints contain many distortions produced by DIBR technology, especially the geometric distortion, which has many different structural characteristics compared with the distortions in natural images [2]. However, the existing Image Quality Assessment (IQA) methods, mainly devised for natural images, are limited in capturing the geometric distortions that occur in DIBR-synthesized images. With the consideration, it is necessary to design an effective and efficient objective assessment metric to judge the DIBR-synthesized images.

In the past, a substantial number of IQA metrics have been proposed for natural images. According to the amount of reference information needed, IQA models can be divided into three types, which include Full-Reference (FR), Reduce-Reference (RR) and No-Reference (NR) or blind metrics. Because the assessment methods, designed for the natural images, have difficulty in capturing the artifacts in DIBR-synthesized images, several recent studies have been proposed to evaluate DIBR-synthesized images. Battisti et $al.$ [3] proposed 3D-SWIM model, which analysed the similarity of statistical features from wavelet subbands between original and distorted DIBR-synthesized images. Sandić-Stanković et $al.$ proposed MW-PSNR [4] and MP-PSNR [5] 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 proposed MP-PSNR-reduce, which achieved better performance and higher efficiency [6]. Yue et $al.$ proposed a DIBR-Synthesized IQA metric based upon combining local and global measures [7]. Vinit et $al.$ designed a RR IQA metric for screen content images and DIBR-synthesized images based on perceptual relevant prediction model, which focus on the importance of textual regions in quality assessment [8]. Gu et $al.$ proposed a natural scene statistics (NSS) model,

This work was supported by the National Natural Science Foundation of China under Grants 61671332,61671336,U1736206, Hubei Province Technological Innovation Major Project under Grant 2017AAA123 and National Key R&D Project under Grant 2016YFE0202300. Corresponding author: Zhongyuan Wang. E-mail: wzy_hope@163.com.

(a) clean image (b) geometric distorted image

Fig. 1: An example of clean and DIBR-synthesized images.

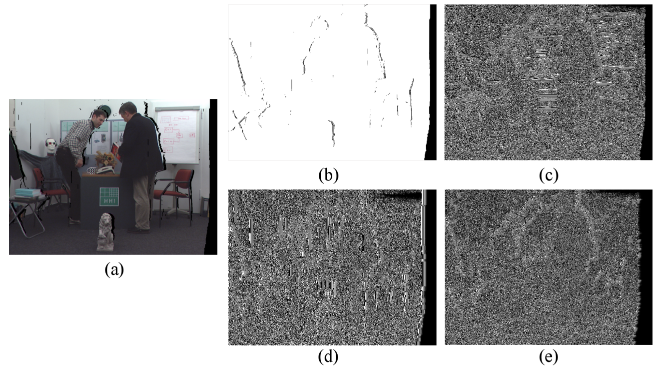


Fig. 2: An example of a DIBR-synthesized image and its corresponding wavelet subbands. (a) DIBR-synthesized image, (b) LL subband, (c) HL subband, (d) LH subband, (e) HH subband.

namely APT [9], which can evaluate the DIBR-synthesized views effectively. In the aforementioned metrics, only [7] and [9] belong to NR IQA metric. However, the performance of [7] is not particularly effective and APT [9] is computationally expensive. Therefore, an real-time reliable blind assessment model is particularly favorite.

In this research, a novel blind quality metric for DIBR-synthesized images is proposed via measuring Geometric Distortions in discrete wavelet transform domain and Image Complexity (GDIC). DIBR-synthesized image is decomposed into wavelet subbands (LL, LH, HL and HH subbands) by using the Cohen-Daubechies-Fauraue 9/7 filter [10]. In LL subband, the high-frequency information of DIBR-synthesized image is removed, so geometric distortions can be exactly detected. Then, we adopt canny operator to capture the edge of wavelet subbands and compute the edge similarity between low-frequency subband (LL) and high-frequency subbands (LH, HL and HH), which denotes the geometric distortions in DIBR-synthesized images. Finally, the quality score is computed by normalizing the captured geometric distortions via image complexity, which is intended for eliminating the influence of image content diversity on image quality assessment. Experiments shows our GDIC is superior to prevailing IQA metrics devised for natural and DIBR-synthesized images.

2. PROPOSED ALGORITHM

Fig. 1 shows that geometric distortions have a remarkable destructive effect on naturalness attribute of an DIBR-synthesized image. Compared to clean image, the human visual system can easily observe geometric distortions, which are marked by red boxes. The geometric distortion badly destroys the quality of DIBR-synthesized image. In the research, a blind quality assessment metric for DIBR-synthesized images is proposed via measuring geometric distortions and image complexity.

Our proposed metric contains two parts, including geometric distortion evaluation and image complexity estimation. The image complexity is used to eliminate the influence of image content diversity on IQA metrics. This is also the difficulty of designing NR metrics relative to FR and RR models.

2.1. Geometric Distortion Detection and Evaluation

The key problem in evaluating DIBR-synthesized images effectively is whether the metric can accurately detect and quantify geometric distortions. Since DIBR-synthesized images can be decomposed into low and high frequency subbands and the high frequency information can be eliminated in the low frequency subband. so, the shape of geometric distortion can be detected exactly by adopting the edge operator.

In this work, the DIBR-synthesized image is decomposed into low-frequency subband and high-frequency subbands by adopting the Cohen-Daubechies-Fauraue 9/7 filter. Fig. 2 shows an example of a DIBR-synthesized image and its corresponding wavelet subbands. Obviously, we can detect the geometric distortion easily and exactly from LL subband via adopting edge operator. In our work, the canny operator is used to detect the edge of decomposed wavelet subbands. The edge detection results of decomposed wavelet subbands in Fig. 2 are shown in Fig. 3. The C_{LL}, C_{HL}, C_{LH} and C_{HH} denote the results that the LL, HL, LH and HH subbands are detected by canny operator.

The edge structural similarity between C_{LL} and C_{HL}, C_{LH} and C_{HH} are used to quantify geometric distortion. The method of calculation is as follows:

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

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

$$S_D = \frac{1}{L}\sum_{l=1}^{L}\left(\frac{2C_{LL}(l)\cdot C_{HH}(l)+\varepsilon}{C_{LL}(l)^2+C_{HH}(l)^2+\varepsilon}\right), \tag{3}$$

where ε is a constant number 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 distortions from horizontal, vertical and diagonal directions.

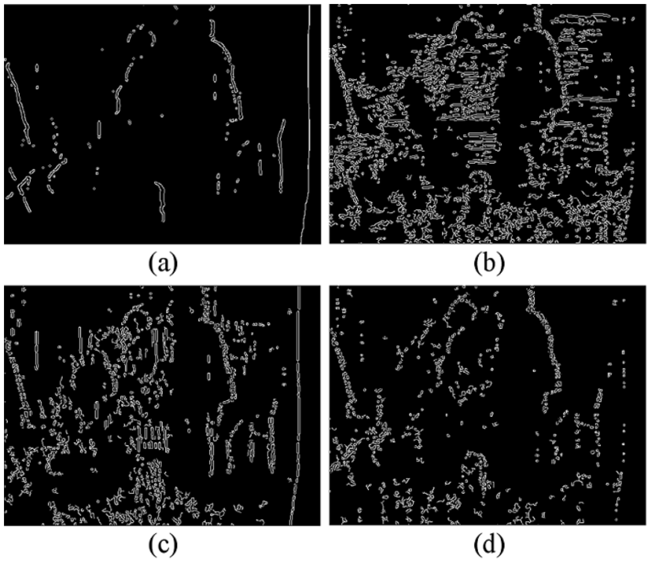


Fig. 3: Edge detection results of decomposed wavelet subbands in Fig. 2,(a) edges of LL subband, (b) edges of HL subband, (c) edges of LH subband, (d) edges of HH subband.

This is also closer to the multi-directional character of the human visual system.

2.2. Image Complexity Estimation

Image complexity is an essential concept in human basic perception to visual stimulus. Image complexity is an key factor to be considered when designing DIBR-synthesized image quality assessment 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 compare to high-complexity, since textures and edges are more difficult to self-described than smooth regions. In [11], [12] and [13], Gu $et\ al.$ constructed an novel hybrid filter, systematically combining the Autoregressive (AR) and bilateral (BL) filter, which can be used to estimate the image complexity. The expression of the hybrid filter is as follows:

$$\hat{y}_q = \frac{Q^n(x_q)\hat{a} + kQ^n(x_q)b}{1+k}, \tag{4}$$

where x_q is the pixel location in a given DIBR-synthesized image; $Q^n(x_q)$ is composed of the n neighboring pixels of x_q; $k = 9$ adjusts the relative strength of the responses of the AR and BL filters; $\hat{a}$ and b represent the coefficients of AR and BL filter respectively. The detailed solution of the coefficients $\hat{a}$, b and k can be seen in the article [11]. The image complexity is estimated as follows:

$$F = -\int H'_{(\rho)} \log H'_{(\rho)} d_\rho, \tag{5}$$

where $H'_{(\rho)}$ represents the probability density of grayscale ρ 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.

2.3. Proposed DIBR-synthesized Image Quality Metric

In order to eliminate the effect of image content diversity on IQA metric, the image complexity is used to normalize the quantized geometric distortions, so the overall quality score Q is computed as:

$$Q = \frac{\sum \omega_i S_i}{F}, \tag{6}$$

where $i = H, V$ and D; Adjusting the weight coefficient of ω_i can improve the performance of the quality evaluation metric; The ω_i are simply set to 1 in this paper. The higher quality DIBR-synthesized images have less geometric distortions. Therefore, lower Q value shows the better image quality.

3. EXPERIMENTAL RESULTS

3.1. Experimental Settings

In this research, we design a valid blind quality assessment metric, which specifically matches the characteristic of the DIBR-synthesized images. To test the performance of our proposed GDIC, the IRCCyN/IVC database [2] is used to check the effectiveness of the proposed metric. The database contains 12 original images and their associated 84 synthesized views, corrupted by geometric distortions, which were processed via 7 different DIBR algorithms.

Three diffusely employed criteria, namely Pearson Linear Correlation Coefficient (PLCC), Spearman Rank order Correlation Coefficient (SRCC) and Root Mean Square Error (RMSE), are adopted to evaluate the performance of IQA metrics. PLCC, RMSE and SRCC are employed to measure the prediction accuracy and estimate the prediction monotonicity respectively. A better IQA metric should obtain a higher value of PLCC and SRCC, while achieve a lower value of RMSE. They are calculated following a five-parameter nonlinear mapping:

$$f(x) = \tau_1\left(\frac{1}{2} - \frac{1}{1+e^{\tau_2(x-\tau_3)}}\right) + \tau_4 x + \tau_5, \tag{7}$$

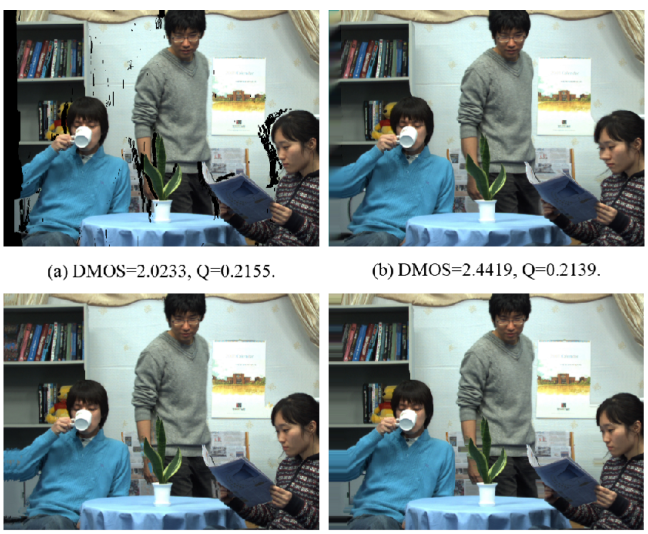
where $\tau_i, i = 1,2,3,4,5$ are the parameters to be fitted; x and $f(x)$ represent the predicted score and its associated subjective score.

3.2. Performance Evaluation

We have compared our proposed metric with two types of IQA models. The first type of IQA metric was designed for natural images. It include FR IQA metrics (such as PSNR,

Table 1: Performance comparison between state-of-the-art models. The best performance in each type is highlighted.

Metric	PLCC	SRCC	RMSE	Designed for	Category
PSNR	0.3976	0.3095	0.6109	Natural Images	Full-Reference
SSIM [14]	0.4850	0.4368	0.5823	Natural Images	Full-Reference
VSNR [15]	0.4370	0.3851	0.5989	Natural Images	Full-Reference
IW-SSIM[16]	**0.5831**	0.4053	**0.5409**	Natural Images	Full-Reference
FSIM [17]	0.5828	0.4148	0.5411	Natural Images	Full-Reference
PSIM [18]	0.5315	0.4576	0.5640	Natural Images	Full-Reference
NIQE[19]	0.4374	0.3739	0.5987	Natural Images	No-Reference
QAC [20]	0.3519	0.3108	0.6232	Natural Images	No-Reference
SIQE [11]	0.3219	0.0739	0.6304	Natural Images	No-Reference
ILNIQE [21]	0.4998	**0.5348**	0.5767	Natural Images	No-Reference
3D-SWIM [3]	0.6584	0.6156	0.5011	DIBR-Synthesized Images	Full-Reference
MW-PSNR [4]	0.5622	0.5757	0.5506	DIBR-Synthesized Images	Full-Reference
MP-PSNR [5]	0.6174	0.6227	0.5238	DIBR-Synthesized Images	Full-Reference
MP-PSNR-reduce [6]	0.6772	0.6634	0.4899	DIBR-Synthesized Images	Reduce-Reference
Vinit [8]	0.7145	0.6293	0.4659	DIBR-Synthesized Images	Reduce-Reference
Yue [7]	0.6750	0.6520	0.4620	DIBR-Synthesized Images	No-Reference
APT [9]	0.7307	0.7157	0.4546	DIBR-Synthesized Images	No-Reference
Proposed GDIC	**0.7332**	**0.7551**	**0.4528**	DIBR-Synthesized Images	No-Reference



(a) DMOS=2.0233, Q=0.2155.

(b) DMOS=2.4419, Q=0.2139.

(c) DMOS=2.9070, Q=0.2099.

(d) DMOS=2.9535, Q=0.2091.

Fig. 4: Subjective and objective scores (predicted by proposed GDIC) of four DIBR-synthesized images with different distortion levels.

SSIM [14], VSNR [15], IW-SSIM [16], FSIM [17] and PSIM [18]) and NR IQA metrics (such as NIQE [19], QAC [20], SIQE [11] and ILNIQE [21]). The second type of IQA metric was devised for DIBR-synthesized images. It contain FR IQA algorithms (such as 3D-SWIM [3], MW-PSNR [4] and MP-PSNR [5]), RR IQA metrics (MP-PSNR-reduce [6] and Vinit [8]) and NR IQA metrics (Yue [7] and APT [9]).

The experimental results of the aforementioned IQA metrics are shown in Table 1. The results indicate the IQA metrics for natural images are limited in evaluating DIBR-synthesized images. The best PLCC (0.5831), RMSE (0.5409) and SRCC (0.5348) are obtained by IW-SSIM and ILNIQE respectively. Existing DIBR-synthesized image quality metrics produce better results and the highest values of PLCC (0.7307), SRCC (0.7157) and RMSE (0.4546) obtained by APT. Compared to other metrics, the proposed GDIC achieves the best performance and the values of PLCC, SRCC and RMSE are 0.7332, 0.7551 and 0.4528 respectively. Fig. 4 show four DIBR-synthesized images with different distortion levels. Obviously, with the increase of subjective scores (DMOS), the objective scores (Q) predicted by our GDIC decreases gradually. So our GDIC are highly consistent with subjective ratings for DIBR-synthesized views.

4. CONCLUSION

In this research, we have proposed a novel blind quality assessment metric for DIBR-synthesized images. The metric is composed of two parts, including geometric distortions quantization and image complexity estimation. The first part is conducted to detect and quantify the geometric distortions. In the second part, image complexity is used to eliminate the impact of image content diversity on IQA metrics. Experiments conducted on IRCCyN/IVC database show the superiority of our blind quality algorithm as compared with prevailing existing FR, RR and NR methods, which refer to as two types of IQA metrics intended for natural and 3D-synthesized images.

5. REFERENCES

[1] Y. Yang, Q. Liu, X. He, and Z. Liu, "Cross-view multilateral filter for compressed multi-view depth video," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 302–315, Jan 2019.

[2] E. Bosc, R. Pépion, and P.L. Callet, "Towards a new quality metric for 3-d synthesized view assessment," *IEEE J. Select. Top. Signal Process.*, vol. 5, no. 7, pp. 1332–1343, Nov 2011.

[3] 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, 2015.

[4] D. Sandic-Stankovic, D. Kukolj, and P. Le Callet, "Dibr-synthesized image quality assessment based on morphological wavelets," *in Proc. IEEE Int. Workshop Quality Multimedia Experience*, 2015.

[5] D. Sandic-Stankovic, D. Kukolj, and P. Le Callet, "Dibr-synthesized image quality assessment based on morphological multi-scale approach," *True Vision-Capture, Transmiss. Display 3D Video*, vol. 1, 2015.

[6] D. Sandic-Stankovic, 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.

[7] 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.*, 2018.

[8] V. Jakhetiya, K. Gu, W. Lin, Q. Li, and S.P. Jaiswal, "A prediction backed model for quality assessment of screen content and 3d synthesized images," *IEEE Trans. Ind Inform.*, vol. 14, no. 2, pp. 652–660, Feb 2018.

[9] 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.

[10] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Commun. Pur. Appl. Math.*, vol. 45, no. 5, pp. 485–560, 1992.

[11] K. Gu, J. Zhou, J.F. Qiao, G.T. Zhai, and A.C. Bovik W.S. Lin, "No-reference quality assessment of screen content pictures," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 4005–4018, Aug 2017.

[12] 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. Cybernetics.*, vol. 47, no. 12, pp. 4559–4565, Dec 2017.

[13] K. Gu, J.F. Qiao, X.K. Min, G.H. Yue, and D. Thalmann W.S. Lin, "Evaluating quality of screen content images via structural variation analysis," *IEEE Trans. Visual. Comput. Graphics*, vol. 24, no. 10, pp. 2689–2701, 2018.

[14] 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.

[15] D.M. Chandler and S.S. Hemami, "Vsnr: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2284–2298, Sep 2007.

[16] 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.

[17] 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.

[18] 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.

[19] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Process. Lett.*, vol. 22, no. 3, pp. 209–212, Mar 2013.

[20] W. Xue, L. Zhang, and X. Mou, "Learning without human scores for blind image quality assessment," *IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2013.

[21] 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.