QUALITY ASSESSMENT OF MULTI-VIEW-PLUS-DEPTH IMAGES

Jiheng Wang[1], *Shiqi Wang*[2], *Kai Zeng*[1] *and Zhou Wang*[1]

[1]Dept. of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada
[2]Rapid-Rich Object Search Laboratory, Nanyang Technological University, Singapore

ABSTRACT

Multi-view-plus-depth (MVD) representation has gained significant attention recently as a means to encode 3D scenes, allowing for intermediate views to be synthesized on-the-fly at the display site through depth-image-based-rendering (DIBR). Automatic quality assessment of MVD images/videos is critical for the optimal design of MVD image/video coding and transmission schemes. Most existing image quality assessment (IQA) and video quality assessment (VQA) methods are applicable only after the DIBR process. Such post-DIBR measures are valuable in assessing the overall system performance, but are difficult to be directly employed in the encoder optimization process in MVD image/video coding. Here we make one of the first attempts to develop a perceptual pre-DIBR IQA approach for MVD images by employing an information content weighted approach that balances between local quality measures of texture and depth images. Experiment results show that the proposed approach achieves competitive performance when compared with state-of-the-art IQA algorithms applied post-DIBR.

Index Terms— 3D image, image quality assessment, depth-image-based-rendering, multi-view-plus-depth

1. INTRODUCTION

With the fast development of 3D acquisition, communication, processing and display technologies, automatic quality assessment of 3D images and videos has become ever important. Many advanced 3D video systems are based on multi-view-plus-depth (MVD) representation [1], where typically two or three views of texture and depth videos are encoded. This allows for intermediate views to be synthesized on-the-fly at the display site from the decoded texture and depth views by means of depth-image-based-rendering (DIBR) [2]. Objective quality assessment of 3D synthesized images/videos from DIBR is a challenging problem [3]. Zhang *et al.* proposed a full reference synthesized video quality method by combining the spatial and temporal distortions of the synthesized view [4]. Liu *et al.* conducted a subjective study of synthesized single-view videos with texture and depth compression and proposed an objective model emphasizing on temporal flicker distortion induced by depth compression and view synthesis processes [5]. These existing studies are valuable but limited in one common aspect, i.e., the quality assessment process involves view synthesis and is applicable only after the DIBR process. The drawback of such methods, which we call post-DIBR approaches, is that they are difficult to be directly employed in the design and optimization of MVD based 3D video coding systems, because the actual texture and depth video signals being encoded and transmitted are the views before DIBR.

In the literature, only a few objective models predict the quality of synthesized views before the DIBR process. Jang *et al.* proposed a fast quality metric purely for depth maps without view synthesis [6]. However, neither the impact of texture distortions nor the interaction between texture and depth distortions was taken into consideration. Fang *et al.* proposed an analytical model to estimate the quality of synthesized views [7], where texture-induced errors and depth-induced errors are combined for synthesis distortion estimation in terms of the mean square error (MSE) between the synthesized views from the original and decoded texture images and depth maps. Oh *et al.* defined a rendering view distortion function and presented an efficient depth map coding scheme [8]. Similarly, this MSE-based rendering view distortion is determined by texture information and depth coding errors together, and the relationship is multiplicative. The limitation of such MSE-based pre-DIBR methods is that they do not predict image/video quality from the perceptive of visual perception, and thus are not useful in the perceptual MVD encoder optimization process.

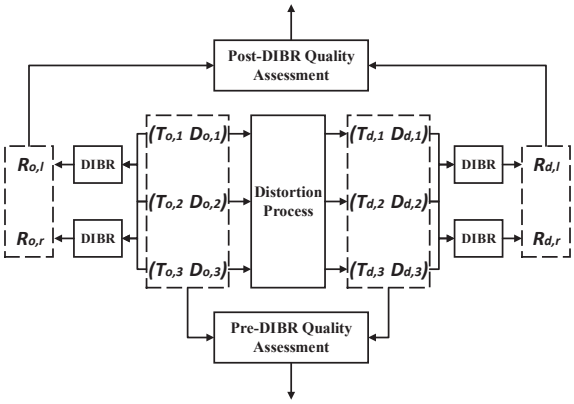


Fig. 1: Two types of quality assessment of stereoscopic 3D images rendered from MVD representations. $(T_{o,1},\ D_{o,1})$, $(T_{o,2},\ D_{o,2})$, $(T_{o,3}, D_{o,3})$ and $(T_{d,1},\ D_{d,1})$, $(T_{d,2},\ D_{d,2})$, $(T_{d,3},\ D_{d,3})$ are the (texture, depth) images for 3 views in the original and distorted MVD representations, respectively. $(R_{o,l},\ R_{o,r})$ and $(R_{d,l},\ R_{d,r})$ are the (left, right) views of DIBR-synthesized original and distorted stereoscopic 3D images, respectively. Type 1: post-DIBR 3D-IQA; Type 2: pre-DIBR 3D-IQA.

The main purpose of this work is to develop a perceptual pre-DIBR image quality assessment (IQA) approach for MVD images, aiming for establishing a more convenient IQA model that can be used in the design of MVD coding schemes. A diagram that explains the difference between the two types (pre-DIBR and post-DIBR) of MVD quality assessment problems, as well as how multiple texture and depth images are used to synthesize stereoscopic 3D images are shown in Fig. 1. In fact, in addition to all existing stereoscopic 3D-IQA models, any 2D-IQA method may also be ap-

plied for post-DIBR quality assessment by averaging 2D-IQA results of stereoscopic views. But none of them has been shown to be useful in the pre-DIBR case. Our work starts by observing how human subjects evaluate synthesized MVD images, how post-DIBR methods perform in predicting subjective quality, and how the performance varies depending on distortion types. Such observations help us develop our pre-DIBR algorithm, which demonstrates competitive performance against post-DIBR approaches.

2. POST-DIBR QUALITY ASSESSMENT

Subjective testing is critical in understanding IQA problems and validating IQA models. A highly valuable subjective study was introduced in [9], which resulted in an MCL-3D database for 3D-IQA using 2D-image-plus-depth source. The database was created from nine pristine image-plus-depth source contents. The resolution of texture and depth images is 1920×1080 or 1024×768. Each texture and depth image was altered by six types of distortions: additive white Gaussian noise contamination, Gaussian blur, downsampling blur, JPEG compression, JP2K compression and transmission loss. Each distortion type had four distortion levels. Three types of rendering combinations, i.e., distorted texture images and original depth maps (Texture-Distortion-Only), original texture images and distorted depth maps (Depth-Distortion-Only), and distorted texture images and distorted depth maps (Texture-Depth-Distortion), are used as the input into the DIBR software to render the distorted stereopairs. In total, there are 657 rendered stereoscopic images (including 9 "original" stereopairs). Fig. 1 illustrates the building process of MCL-3D database. Pair-wise comparison was adopted in the subjective test and the mean opinion score (MOS) was computed for each distorted rendered stereopair. More detailed descriptions of this database and the subjective experiment can be found in [9].

Post-DIBR methods directly work on synthesized stereoscopic pairs. Let ($R_{o,l}$, $R_{o,r}$) and ($R_{d,l}$, $R_{d,r}$) be the left- and right-view image pairs of the rendered stereoscopic images from the original and distorted texture images and depth maps, respectively. For the post-DIBR case, we are interested in investigating the appropriateness of existing 2D/3D-IQA methods to predict 3D quality of synthesized views by comparing ($R_{o,l}$, $R_{o,r}$) and ($R_{d,l}$, $R_{d,r}$).

We first test 2D-IQA methods, which can be applied to the left- and right-view images independently and then averaged to predict 3D quality. Previous studies suggested that in the case of symmetric distortion of both views (in terms of both distortion types and levels), simply averaging state-of-the-art 2D-IQA measures of both views is sufficient to provide reasonably accurate quality predictions of stereoscopic images [10, 11] and stereoscopic videos [12].

Note that given a pair of distorted texture image and depth map, multi-view rendering and coding schemes usually generate rendered views at a similar distortion level. In the MCL-3D database, for each pair of texture image and depth map, there are two synthesized views rendered from the view synthesis reference software (VSRS) [13]. In addition, we employed another commonly used rendering software, i.e., 3D-HTM [14], to generate six different rendered views for each pair of texture image and depth map. In total, there are 2×657 (9 "original" and 648 distorted) rendered views from VSRS and 6×657 rendered views from 3D-HTM, respectively. We then compare the perceptual 2D quality of different rendered views within VSRS method, within 3D-HTM method, and between VSRS and 3D-HTM methods using full reference (FR) 2D-IQA methods including PSNR, SSIM [15], MS-SSIM [16], and IW-SSIM [17] and report the average Pearson's linear correlation coefficient (PLCC), Spearman's rank-order correlation coefficient (SRCC) and Root Mean Squared Error (RMSE) scores between different rendered views in Table 1. Higher PLCC and SRCC and lower RMSE indicate better consistency in terms of the perceptual quality between different rendered views. From Table 1, it can be observed that the perceptual quality of rendered views is generally indistinguishable within VSRS or 3D-HTM, and the performance drop in the case of mixed VSRS and 3D-HTM views is only moderate. These results suggest that the perceptual quality of rendered views is more affected by the input texture images and depth maps, and the impact of the rendering algorithms is relatively minor. As a result, pre-DIBR IQA becomes a feasible and practically useful approach in MVD image/video applications.

Table 1: Performance Comparison of 2D-IQA models on the MCL-3D database for different rendered views

	2D-IQA	PLCC	SRCC	RMSE
Within VSRS	PSNR	0.9874	0.9892	1.1710
	SSIM	0.9906	0.9928	0.0142
	MS-SSIM	0.9926	0.9938	0.0129
	IW-SSIM	0.9930	0.9934	0.0163
	2D-IQA	PLCC	SRCC	RMSE
Within 3D-HTM	PSNR	0.9917	0.9922	1.0370
	SSIM	0.9957	0.9946	0.0108
	MS-SSIM	0.9969	0.9957	0.0092
	IW-SSIM	0.9945	0.9938	0.0152
	2D-IQA	PLCC	SRCC	RMSE
Between VSRS and 3D-HTM	PSNR	0.9390	0.9541	2.9455
	SSIM	0.9671	0.9668	0.0352
	MS-SSIM	0.9736	0.9710	0.0328
	IW-SSIM	0.9639	0.9664	0.0439

Table 2: Performance Comparison of 2D/3D-IQA models on the MCL-3D database

Category		Method	PLCC	SRCC	RMSE
Post-DIBR	FR 2D-IQA	PSNR	0.832	0.840	1.444
		SSIM [15]	0.894	0.903	1.168
		MS-SSIM [16]	0.864	0.875	1.308
		IW-SSIM [17]	**0.920**	**0.926**	**1.021**
		FSIM [18]	**0.925**	**0.932**	**0.990**
	NR 2D-IQA	BLIINDS-II [19]	0.505	0.499	2.246
		BRISQUE [20]	0.686	0.647	1.893
		CORNIA [21]	0.740	0.734	1.750
		LPSI [22]	0.503	0.449	2.249
		M_3 [23]	0.590	0.506	2.101
		NIQE [24]	0.743	0.599	1.741
	FR 3D-IQA	Benoit [25]	0.871	0.874	1.277
		Chen [26]	0.881	0.884	1.230
		Lin [27]	0.870	0.869	1.285
		Shao [28]	0.848	0.853	1.380
		Yang [29]	0.835	0.845	1.433
		You [30]	0.892	0.904	1.175
		Zhang [31]	**0.934**	**0.939**	**0.930**
Pre-DIBR		Proposed	**0.911**	**0.916**	**1.074**

The 2D-IQA models being tested include FR methods PSNR, SSIM, MS-SSIM, IW-SSIM, and FSIM [18], and no-reference (NR) methods, BLIINDS-II [19], BRISQUE [20], CORNIA [21], LPSI [22], M_3 [23], and NIQE [24]. Table 2 reports PLCC, SRCC and RMSE results between 3D-MOS scores and the average 2D-IQA measurements. The corresponding scatter plots in terms of IW-SSIM are shown in Fig. 2. From Table 2 and Fig. 2, it can be observed that directly averaging the FR 2D-IQA measures of both views provides accurate image quality predictions of rendered stereopairs, which is consistent with previous findings in [10, 11, 12]. However, if the "reference" stereopairs are not available, there is a large drop in the performance from FR to NR models.

We also test some recent methods that are designed for FR 3D-IQA. The results on MCL-3D database are given in Table 2, where it can be seen that most FR 3D-IQA methods do not show superi-

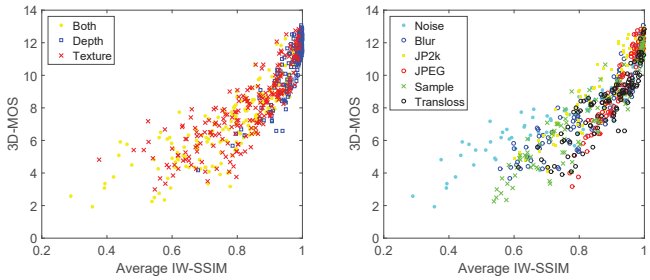


Fig. 2: 3D-MOS versus predictions from IW-SSIM of 2D left- and right-views. (a) By different rendering combinations. (b) By different distortion types.

ority over FR 2D-IQA methods. The best result is obtained by the 3D-MAD method [31], which demonstrates similar performance as compared to IW-SSIM and FSIM, purely 2D algorithms. Another important observation is that all perception-based 2D-IQA methods outperform PSNR, which suggests that the performance of the MSE-based pre-DIBR methods may be improved by incorporating perceptual modelling.

3. PRE-DIBR QUALITY ASSESSMENT

3.1. Distortion type dependency

As mentioned earlier, in the rendering process of MCL-3D database, three types of combinations, i.e., Texture-Distortion-Only, Depth-Distortion-Only, and Texture-Depth-Distortion, were adopted to create different kinds of distorted stereopairs. Fig. 3 shows 3D-MOS scores for different distortion types and different rendering combinations. We compute SSIM between $(R_{o,l}, R_{d,l})$ to measure local quality/distortion of the distorted rendered left-view images. The resultant SSIM index maps for the cases of Texture-Depth-Distortion, Texture-Distortion-Only and Depth-Distortion-Only are denoted as $S^{T+D}_{R,l}$, $S^{T}_{R,l}$ and $S^{D}_{R,l}$, respectively, as shown in Column (a), (b), and (c) in Fig. 4, where brightness indicates the magnitude of the local SSIM index (i.e., brighter = better quality).

From Fig. 3 and Fig. 4, it can be observed that there exists a specific distortion type dependency with respect to 3D-MOS scores for different rendering combinations. For blur, JP2K, JPEG and downsample distortion, 3D-MOS scores for Texture-Depth-Distortion and Texture-Distortion-Only images are very close, which both are strictly decreasing with the increase of distortion level. For Depth-Distortion-Only images, 3D-MOS scores are always at a high quality level and not falling down with the increasing level. Similar observations can be found in Fig. 4. For Texture-Depth-Distortion and Texture-Distortion-Only images at Column (a) and (b), the SSIM maps exhibit almost the same local spatial variations and intensity levels, indicating a very close overall quality for both cases; for Depth-Distortion-Only images, a significantly better overall quality is presented as darker pixels only appear in a very small region. For noise and transmission loss, 3D-MOS scores lie in different quality levels for different rendering combinations with the increasing order of Texture-Depth-Distortion, Texture-Distortion-Only and Depth-Distortion-Only. Also, 3D-MOS scores decrease with the increase of distortion levels in all cases. While in Fig. 4, the SSIM maps for Texture-Depth-Distortion and Texture-Distortion-Only images exhibit similar spatial variations but relative brighter pixels can be seen

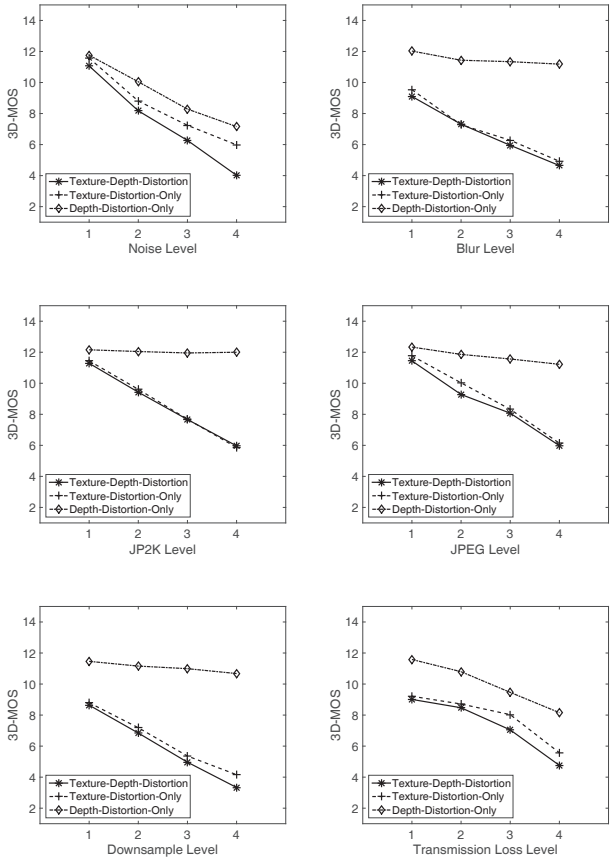


Fig. 3: 3D-MOS scores for different distortion types and different rendering combinations.

in the Texture-Distortion-Only case, indicating a better overall quality; for Depth-Distortion-Only images, an even better overall quality compared to the Texture-Distortion-Only images is pronounced as darker pixels are significantly less visible.

In general, the distortions from noise contamination and transmission loss affect the 3D quality of rendered stereopairs more significantly than from blur, JP2K, JPEG and downsampling. Another important observation is that the impact of texture image distortions are much stronger compared to that of depth map distortions.

3.2. Pre-DIBR 3D-IQA algorithm

Let $(T_{o,1}, D_{o,1})$, $(T_{o,2}, D_{o,2})$, $\cdots$, $(T_{o,M}, D_{o,M})$ be the original texture images and their associated depth maps, and $(T_{d,1}, D_{d,1})$, $(T_{d,2}, D_{d,2}), \cdots, (T_{d,M}, D_{d,M})$ be the corresponding distorted texture images and depth maps for M views. Fig. 1 gives an example for the case of $M = 3$. We propose a top-down model to predict the 3D quality of synthesized views before DIBR. Specially, the prediction of Q^{3D} is calculated by directly averaging each view's quality Q^{3D}_k:

$$Q^{3D} = \frac{1}{M}\sum_{k=1}^{M}\left(Q^{3D}_k\right). \quad (1)$$

The procedure to compute each view's quality Q^{3D}_k is shown in Fig. 5. We define an overall quality map $S_{O,k}$, which is a combina-

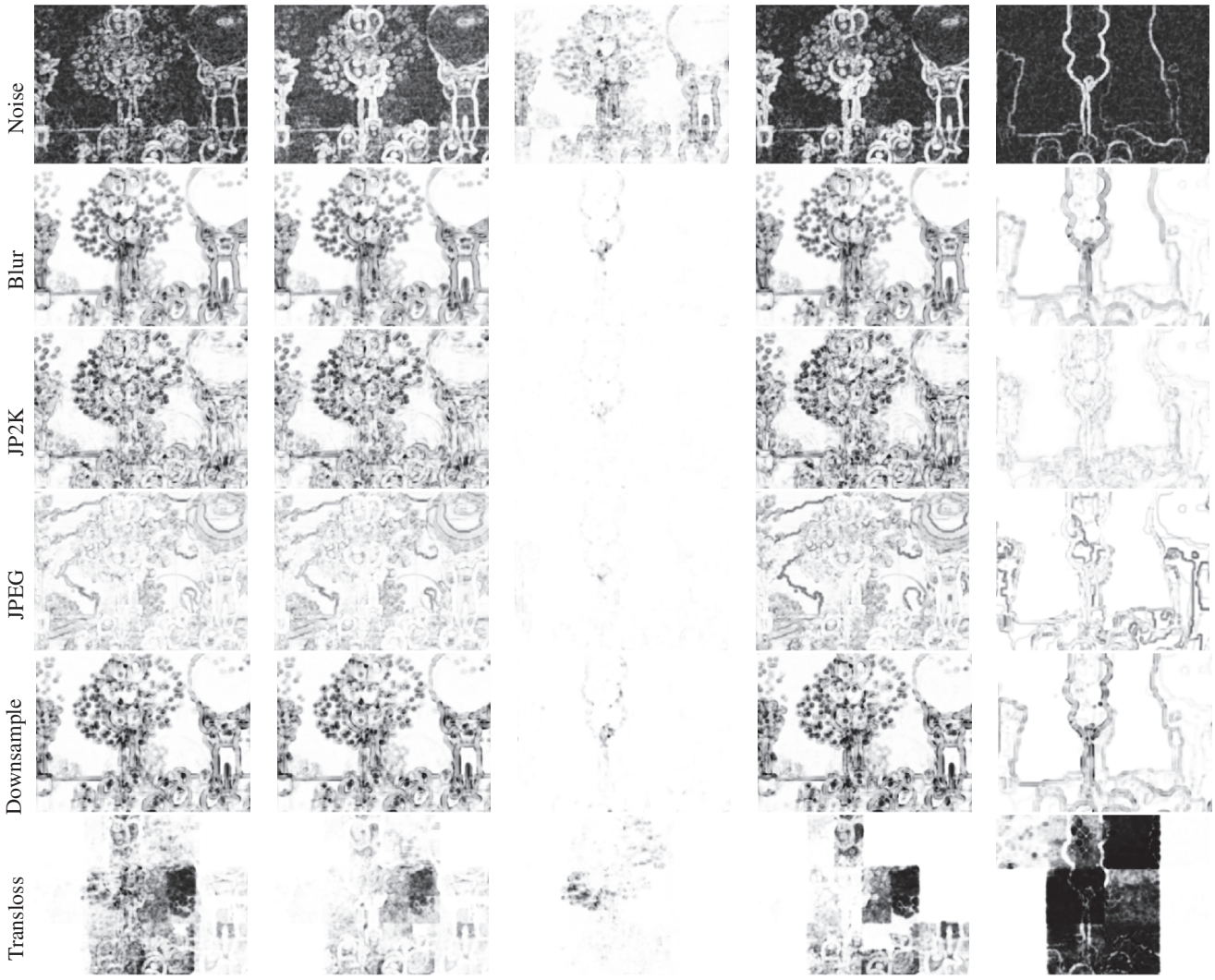


Fig. 4: SSIM maps of the rendered and source Balloon images for all distortion types at level 4. (a): Rendered, Texture-Depth-Distortion; (b): Rendered, Texture-Distortion-Only; (c): Rendered, Depth-Distortion-Only; (d): Source, Texture; (e): Source, Depth.

tion of the k-th view's texture induced quality map $S_{T,k}$ and depth induced quality map $S_{D,k}$:

$$S_{O,k} = w_{T,k} S_{T,k} + w_{D,k} S_{D,k} \,, \tag{2}$$

where $w_{T,k}$ and $w_{D,k}$ are the weights assigned to $S_{T,k}$ and $S_{D,k}$, respectively. Q_k^{3D} is obtained by spatially average pooling over $S_{O,k}$.

The key step is to determine the adaptive weighting factors $w_{T,k}$ and $w_{D,k}$. Here we use an information content weighting approach, where the perceived local information content is quantified as the number of bits that can be received from a statistical image information source that passes through a noisy visual channel. Assume that the source power is P and the channel noise power is C. The mutual information between the source and destination is

$$I = \frac{1}{2} \log \left(1 + \frac{P}{C} \right) . \tag{3}$$

Now assume that the source power of a local image patch can be estimated as $\sigma_{T,k}^2$ and $\sigma_{D,k}^2$ for texture images and depth maps, respectively, then the information maps from the k-th view's texture image and depth map as shown in Fig. 6 are given by

$$I_{T,k} = \log \left(1 + \frac{\sigma_{T,k}^2}{C} \right) \quad \text{and} \quad I_{D,k} = \log \left(1 + \frac{\sigma_{D,k}^2}{C} \right) , \tag{4}$$

where $\sigma_{T,k}^2$ and $\sigma_{D,k}^2$ are the local variance maps by computing local variances at each spatial location. C is set as 0.01 and an 11×11 sliding Gaussian window with standard deviation of 1.5 pixels is employed.

Given the information maps $I_{T,k}$ and $I_{D,k}$, the weights assigned to the k-th texture and depth induced quality maps are given by

$$w_{T,k} = \frac{\sum I_{T,k}}{\sum I_{T,k} + \sum I_{D,k}} \quad \text{and} \quad w_{D,k} = \frac{\sum I_{D,k}}{\sum I_{T,k} + \sum I_{D,k}} , \tag{5}$$

where a spatially average pooling is applied to both $I_{T,k}$ and $I_{D,k}$. The average weights to the texture and depth quality maps are com-

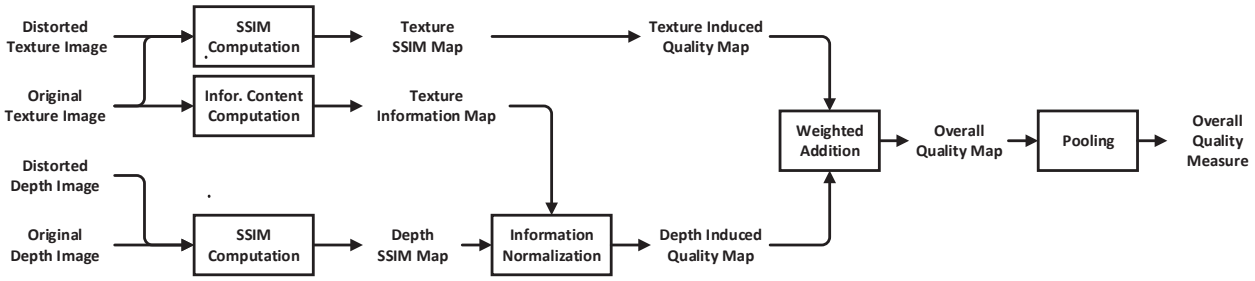
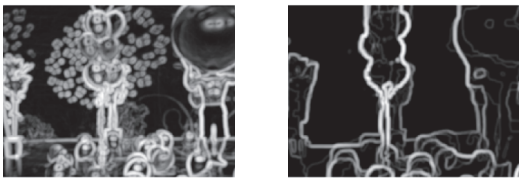


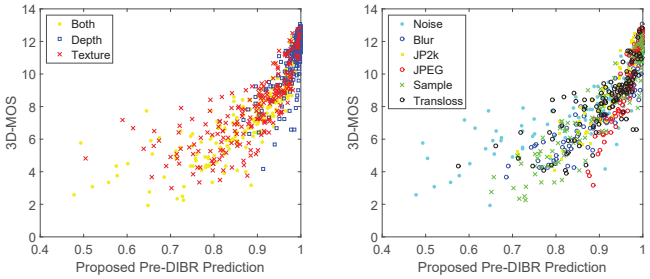
Fig. 5: Diagram of the proposed 2D-to-3D quality prediction model.



(a) Texture Information Map

(b) Depth Information Map

Fig. 6: Information maps of the pristine Balloon texture and depth images.



(a) (b)

Fig. 7: 3D-MOS versus predictions from the proposed Pre-DIBR method. (a) By different rendering combinations. (b) By different distortion types.

puted as 0.7856 and 0.2144, respectively. The standard deviation is 0.0647.

For the k-th original texture image $T_{o,k}$ and distorted texture image $T_{d,k}$, SSIM is used to compute local quality/distortion and the SSIM index map $S_{T,k}$ is obtained. Sample $S_{T,1}$ maps are shown in Column (d) in Fig. 4. For all distortion types, it can be observed that the SSIM maps $S_{T,1}$ and $S^T_{R,l}$ exhibit highly similar local spatial variations. Thus we denote $S_{T,k}$ as the texture induced quality map for the k-th view.

For the k-th original depth map $D_{o,k}$ and distorted depth map $D_{d,k}$, SSIM is again employed to compute local quality/distortion and the SSIM index map $S_{D,k}$ is obtained. Sample $S_{D,1}$ maps are shown in Column (e) in Fig. 4. Unlike texture distortion, SSIM maps $S_{D,1}$ and $S^D_{R,l}$ look significantly different in terms of spatial variations. We found that the impact of depth distortions on the rendered image is more correlated with the information map of the corresponding texture image $I_{T,k}$. As a result, when depth distortion occurs in structural regions in the texture images, the quality of rendered images is more affected. Thus we apply a spatial normalization to $S_{D,k}$ with $I_{T,k}$ and denote the resultant map $S'_{D,k}$ as the depth induced quality map for the k-th view:

$$s'_{D,k,j} = \frac{i_{T,k,j} s_{D,k,j}}{\sum_{j=1}^{N} i_{T,k,j}}, \tag{6}$$

where $s_{D,k,j}$ and $s'_{D,k,j}$ are the local quality values in $S_{D,k}$ and $S'_{D,k}$, respectively, and $i_{T,k,j}$ is the weight assigned to the j-th spatial location (j-th pixel) from $I_{T,k}$. Subsequently, the overall quality map $S_{O,k}$ is combined as the weighted average between $S_{T,k}$ and $S'_{D,k}$ using Eq. (2).

Finally, to emphasize the importance of the information content of texture images, a spatially weighted pooling with the texture information map $I_{T,k}$ is applied to the overall quality map $S_{O,k}$:

$$Q^{3D}_k = \frac{\sum_{j=1}^{N} i_{T,k,j} s_{O,k,j}}{\sum_{j=1}^{N} i_{T,k,j}}, \tag{7}$$

where $s_{O,k,j}$ and $i_{T,k,j}$ are the local quality value in $S_{O,k}$ and the weight assigned to the j-th spatial location from $I_{T,k}$, respectively.

We test the proposed algorithm on the MCL-3D database. The PLCC, SRCC, and RMSE values between 3D-MOS and the predicted Q^{3D} value for all images are given in Table 2 and the corresponding scatter plots are shown in Fig. 7, where we can see that the proposed method performs equally well or better than state-of-the-art FR IQA models applied post-DIBR. Since it is applied pre-DIBR, it has great potentials to be employed in perceptually inspired rate-distortion optimization of MVD video coding systems.

4. CONCLUSIONS

We investigated the problem of objective quality assessment of MVD images, with a main focus on the pre-DIBR case. We found that although existing IQA methods can be applied post-DIBR to provide reasonable quality prediction of MVD images, they are hard to be employed as a guiding criterion in the optimization of MVD video coding and transmission systems. We proposed a novel perceptual pre-DIBR method based on information content weighting of both texture and depth images. Experimental results show that the proposed method demonstrates competitive performance against state-of-the-art IQA models applied post-DIBR. Future work includes incorporating the proposed model in the rate-distortion optimization process of visual perception-based MVD 3D video coding schemes.

5. REFERENCES

[1] K. Müller, P. Merkle, and T. Wiegand, "3-D video representation using depth maps," *Proc. IEEE*, vol. 99, no. 4, pp. 643–656, Apr. 2011.

[2] P. Kauff, N. Atzpadin, C. Fehn, M. Müller, O. Schreer, A. Smolic, and R. Tanger, "Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability," *Signal Process.: Image Commun.*, vol. 22, no. 2, pp. 217–234, Feb. 2007.

[3] E. Bosc, R. Pepion, P. Le Callet, M. Köppel, P. Ndjiki-Nya, M. Pressigout, and L. Morin, "Towards a new quality metric for 3-D synthesized view assessment," *IEEE J. of Sel. Topics in Sig. Process.*, vol. 5, no. 7, pp. 1332–1343, Nov. 2011.

[4] Y. Zhang, X. Yang, X. Liu, Y. Zhang, G. Jiang, and S. Kwong, "High-efficiency 3D depth coding based on perceptual quality of synthesized video," *IEEE Trans. on Image Process.*, vol. 25, no. 12, pp. 5877–5891, Dec. 2016.

[5] X. Liu, Y. Zhang, S. Hu, S. Kwong, C-C J. Kuo, and Q. Peng, "Subjective and objective video quality assessment of 3D synthesized views with texture/depth compression distortion," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4847–4861, Nov. 2015.

[6] W. Jang, T. Chung, J. Sim, and C. Kim, "FDQM: Fast quality metric for depth maps without view synthesis," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 25, no. 7, pp. 1099–1112, July 2015.

[7] L. Fang, N. Cheung, D. Tian, A. Vetro, H. Sun, and O. C. Au, "An analytical model for synthesis distortion estimation in 3D video," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 185–199, Jan. 2014.

[8] B. T. Oh, J. Lee, and D. Park, "Depth map coding based on synthesized view distortion function," *IEEE J. of Sel. Topics in Sig. Process.*, vol. 5, no. 7, pp. 1344–1352, Nov. 2011.

[9] R. Song, H. Ko, and C.-C. Kuo, "MCL-3D: a database for stereoscopic image quality assessment using 2D-image-plus-depth source," *arXiv preprint arXiv:1405.1403*, 2014.

[10] M. Chen, L. K. Cormack, and A.C. Bovik, "No-reference quality assessment of natural stereopairs," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3379–3391, Sept. 2013.

[11] J. Wang, A. Rehman, K. Zeng, S. Wang, and Z. Wang, "Quality prediction of asymmetrically distorted stereoscopic 3D images," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3400 – 3414, Nov. 2015.

[12] J. Wang, S. Wang, and Z. Wang, "Quality prediction of asymmetrically compressed stereoscopic videos," in *Proc. IEEE Int. Conf. Image Process.*, Quebec City, QC, Canada, Sept. 2015, pp. 1 – 5.

[13] M. Tanimoto, T. Fujii, and K. Suzuki, "View synthesis algorithm in view synthesis reference software 3.5 (VSRS3.5) Document M16090, ISO/IEC JTC1/SC29/WG11 (MPEG)," 2009.

[14] "JCT-3V. MV- and 3D-HEVC Reference Software, HTM-14.1.," May, 2015, [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn-3DVCSoftware/tags/HTM-14.1/.

[15] 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.

[16] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," in *Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Nov. 2003, pp. 1398–1402.

[17] 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.

[18] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: a feature similarity index for image quality assessment," *IEEE Trans. on Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[19] 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.

[20] 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.

[21] P. Ye, J. Kumar, L. Kang, and D. Doermann, "Unsupervised feature learning framework for no-reference image quality assessment," in *Proc. IEEE Int. Conf. Computer Vis. and Pattern Recog.*, Providence, RI, USA, June 2012, pp. 1098–1105.

[22] Q. Wu, Z. Wang, and H. Li, "A highly efficient method for blind image quality assessment," in *Proc. IEEE Int. Conf. Image Process.*, Quebec City, Quebec, Canada, Sept. 2015.

[23] W. Xue, X. Mou, L. Zhang, A. C. Bovik, and X. Feng, "Blind image quality assessment using joint statistics of gradient magnitude and laplacian features," *IEEE Trans. Image Process.*, vol. 23, no. 11, pp. 4850–4862, Nov. 2014.

[24] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209–212, Mar. 2013.

[25] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images," *EURASIP J. Image Video Process.*, vol. 2008, pp. 659024, Oct. 2008.

[26] M. Chen, C. Su, D. Kwon, L. K. Cormack, and A. C Bovik, "Full-reference quality assessment of stereopairs accounting for rivalry," *Signal Process.: Image Commun.*, vol. 28, no. 9, pp. 1143–1155, Oct. 2013.

[27] Y. Lin and J. Wu, "Quality assessment of stereoscopic 3D image compression by binocular integration behaviors," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1527 – 1542, Apr. 2014.

[28] F. Shao, W. Lin, S. Gu, G. Jiang, and T. Srikanthan, "Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1940–1953, May 2013.

[29] J. Yang, C. Hou, Y. Zhou, Z. Zhang, and J. Guo, "Objective quality assessment method of stereo images," in *Proc. 3DTV Conf., True Vis.-Capture, Transmiss. Display 3D Video*, Potsdam, Gernamy, May 2009, pp. 1–4.

[30] 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. Int. Workshop Video Process. Quality Metrics Consum. Electron*, Scottsdale, AZ, USA, Jan. 2010, pp. 61–66.

[31] Y. Zhang and D. M. Chandler, "3D-MAD: A full reference stereoscopic image quality estimator based on binocular lightness and contrast perception," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3810–3825, Nov. 2015.