

QUALITY ASSESSMENT OF MULTI-VIEW-PLUS-DEPTH IMAGES

Jiheng Wang¹, Shiqi Wang², Kai Zeng¹ and Zhou Wang¹

¹Dept. of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada

²Rapid-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 syn-

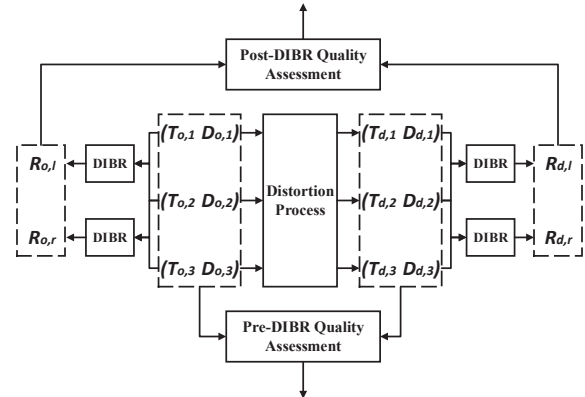


Fig. 1: Two types of quality assessment of stereoscopic 3D images rendered from MVD representations. $(T_{o,1}, D_{o,1})$, $(T_{o,2}, D_{o,2})$, $(T_{o,3}, D_{o,3})$ and $(T_{d,1}, D_{d,1})$, $(T_{d,2}, D_{d,2})$, $(T_{d,3}, D_{d,3})$ are the (texture, depth) images for 3 views in the original and distorted 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.

thesis [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 perspective of visual perception, and thus are not useful in the perceptual MVD encoder optimization process.

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)

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

Within VSRS	2D-IQA	PLCC	SRCC	RMSE
	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
Within 3D-HTM	2D-IQA	PLCC	SRCC	RMSE
	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
Between VSRS and 3D-HTM	2D-IQA	PLCC	SRCC	RMSE
	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

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.

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], M₃ [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 superior

Download English Version:

<https://daneshyari.com/en/article/4991218>

Download Persian Version:

<https://daneshyari.com/article/4991218>

[Daneshyari.com](https://daneshyari.com)