3D-HEVC Visual Quality Assessment: Database and Bitstream Model

Wei Zhou, Ning Liao, Zhibo Chen, and Weiping Li
CAS Key Laboratory of Technology in Geo-spatial Information Processing and Application System
University of Science and Technology of China, Hefei 230027, China
Email: chenzhibo@ustc.edu.cn

***Abstract*—Visual Quality Assessment of 3D/stereoscopic video (3D VQA) is significant for both quality monitoring and optimization of the existing 3D video services. In this paper, we build a 3D video database based on the latest 3D-HEVC video coding standard, to investigate the relationship among video quality, depth quality, and overall quality of experience (QoE) of 3D/stereoscopic video. We also analyze the pivotal factors to the video and depth qualities. Moreover, we develop a No-Reference 3D-HEVC bitstream-level objective video quality assessment model, which utilizes the key features extracted from the 3D video bitstreams to assess the perceived quality of the stereoscopic video. The model is verified to be effective on our database as compared with widely used 2D Full-Reference quality metrics as well as a state-of-the-art 3D FR pixel-level video quality metric.**

Keywords*—*stereoscopic video; database; visual quality assessment; bitstream-level model; 3D-HEVC

I. Introduction

With the rapid development of 3D video technologies and applications, the research on Visual Quality Assessment of 3D/stereoscopic video (3D VQA) has attracted wide interest in both academia and industry. 3D VQA is important for quality monitoring and optimization of stereoscopic video services, which involves in several aspects e.g., multiview and depth compression, transmission, and stereoscopic display. The Joint Collaborative Team on 3D Video Coding Extension Development (JCT-3V) accomplished an advanced 3D extension of High Efficiency Video Coding standard (3D-HEVC) to support efficient representation of multiview video and depth-based 3D video formats [1] in 2015. Therefore, it is desirable to establish a 3D-HEVC video database for investigating the influential factors to the subjective quality of 3D-HEVC compressed video. Moreover, the database with subjective quality labels provides ground truth for developing 3D-HEVC bitstream-level objective video quality assessment models, which are highly efficient for monitoring video quality in networks [2].

Visual quality assessment of 3D video is a challenging problem [3], which involves multiple quality dimensions. Three basic quality dimensions, namely video quality, depth quality, visual discomfort, are identified in [4] to collectively affect the quality of experience (QoE) of 3D video. Some subjective study of these quality aspects have been done in literatures. Recent subjective study based on stereoscopic image databases [5]–[7] showed that in case of symmetric distortion of left and right views, the perceived quality of the binocular combination was approximately the average quality of both views. Whereas, in case of asymmetric distortion, the influence of both views on the binocular combination quality is distortion-type dependent [8] due to the complex human binocular perception mechanism. This phenomenon is also observed for 3D video [9]. Further, the relationship of eye-strain, perceived depth, and perceived image quality is discussed in [10] for stereoscopic image and in [11] for 3D video. However, the 3D video contents in [10] were compressed by H.264 in Intra mode, which is less practical from the perspective of building objective 3D video quality assessment model.

For the convenience of research on objective quality assessment model of stereoscopic video, three public accessible stereoscopic video databases [12]–[14] have been created so far as we know. MMSPG 3D video database [14] only considered the influence of camera baseline parameter on the QoE of 3D video. Different types of symmetric distortions including H.264 and JPEG 2000 compression artifacts, blurriness, and sharpening effects, were created in NAMA3DS1-COSPAD1 database [13]. Both symmetric and asymmetric compression degradations produced with H.264/AVC and HEVC encoders were included in the database of [12]. Note that the distortions in both views of stereoscopic video are manipulated independently. However, in future practice, the multiple views of 3D video will be more likely encoded by exploiting the inter-view correlation as done in 3D-HEVC, and the influence of compression on the optimization of the synthesized view should also be considered.

In this paper, we build a 3D-HEVC compressed video database to study the influence of varying depth range and 3D-HEVC compression artifacts on the visual quality of various 3D video contents. Our ultimate goal of building this database is to develop a bitstream-level 3D-HEVC video quality assessment model. Thus, we further propose a preliminary No-Reference (NR) objective video quality assessment algorithm based on some relevant features extracted from the 3D-HEVC bit-streams. The experiment results demonstrate that our proposed bitstream-level 3D video quality assessment model achieves significant improvement over some widely used Full-Reference (FR) 2D and 3D pixel-level quality assessment algorithms.

For the other parts of the paper, we describe the details of building the 3D-HEVC video database in Section II. In Section III, we discuss the relationship of video quality, depth quality and overall QoE, and investigate the key factors to these qualities. In Section IV, we propose a simple but effective bitstream-level quality assessment model and analyze its performance. Finally, our conclusions and future work plan are presented in Section V.

II. Subjective Experiment Method

To build our 3D-HEVC video database, we select the 3D video content covering various spatial and temporal complexities for the texture video and depth video. We also design different distortion levels of symmetric and asymmetric 3D-HEVC compression.

A. Selection of Video Content

The selected video content covers a wide range of spatial-temporal complexity and disparity complexity. We use the 3D-HEVC testing sequences [15], which are the only publicly available 3D videos of multiview plus depth (MVD) format so far. The disparities of all sequences are within the comfortable viewing zone in our viewing conditions of 3D videos, except the sequence "Newspapers" in the wide baseline case (defined in the next subsection). Thus, we exclude "Newspapers" from our database. Six high-quality pristine stereoscopic videos finally selected in our database are shown in Fig. 1. The spatial and temporal complexities of texture video (SI, TI) and of depth video (DSI, DTI) are calculated according to ITU-T P.910 and shown in Fig. 2. All sequences are of 1920x1088 pixels per frame, except "Balloons" of 1024x768 pixels.

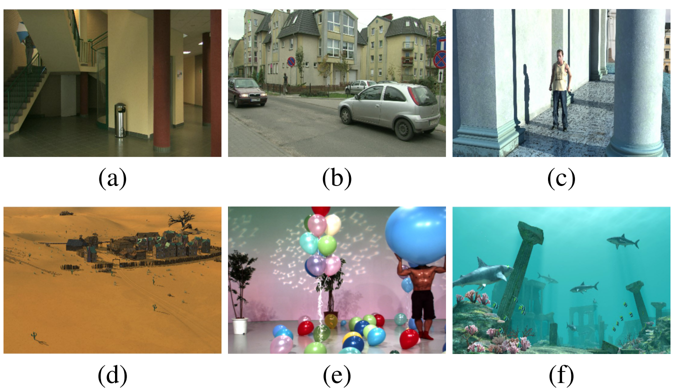

Fig. 1. Snapshots of 6 pristine 3D videos used to create the 3D-HEVC video database. (a) Hall2, (b) Street, (c) Dancer, (d) GTFly, (e) Balloons, (f) Shark

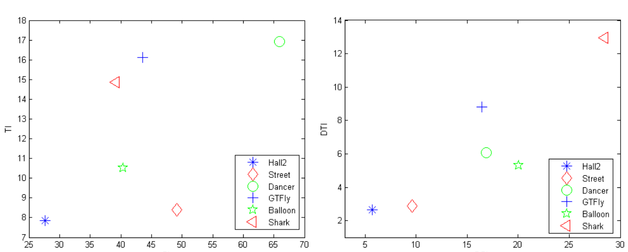

Fig. 2. Left drawing is the SI and TI of texture video; right drawing is the DSI and DTI of depth video

B. Design of Processing Conditions

In this subjective study, we are interested in (1) the influence of symmetric and asymmetric 3D-HEVC compression on the quality of binocular perception, (2) the visibility of 3D-HEVC compression artifacts in different presence of depth, and, (3) the relationship between video quality, depth quality and overall experience of stereoscopic video quality. As a result, we design the Hypothetic Reference Circuits (HRCs) as shown in Table I, to create the Processed Video Sequences (PVS). HRC and PVS are terminologies used in the VQEG [16].

TABLE I. Processing Conditions (HRCs)

Num	HRCID	QP,base-view /dependent-view	Baseline
1	HRC001	-	2D/Zero baseline
2	HRC002	25,25	2D/Zero baseline
3	HRC003	35,35	2D/Zero baseline
4	HRC004	45,45	2D/Zero baseline
5	HRC005	25,35	2D/Zero baseline
6	HRC006	25,40	2D/Zero baseline
7	HRC007	25,45	2D/Zero baseline
8	HRC008	30,40	2D/Zero baseline
9	HRC101	-	3D,short-baseline
10	HRC102	25 25	3D,short-baseline
11	HRC103	35 35	3D,short-baseline
12	HRC104	45 45	3D,short-baseline
13	HRC105	25 40	3D,short-baseline
14	HRC106	25 50	3D,short-baseline
15	HRC107	30 40	3D,short-baseline
16	HRC108	30 50	3D,short-baseline
17	HRC109	40 50	3D,short-baseline
18	HRC201	-	3D,wide-baseline
19	HRC202	25 25	3D,wide-baseline
20	HRC203	35 35	3D,wide-baseline
21	HRC204	45 45	3D,wide-baseline
22	HRC205	25 40	3D,wide-baseline
23	HRC206	25 50	3D,wide-baseline
24	HRC207	30 40	3D,wide-baseline
25	HRC208	30 50	3D,wide-baseline
26	HRC209	40 50	3D,wide-baseline

We control the camera baseline as follows to obtain videos with different depth: 1) monoscopic (2D) video, where left and right views correspond to the same camera number in the multiview sequence, 2) short-baseline video (3D short baseline), whose 2 views correspond to the recommended 2 views in [15], 3) Wide-baseline video (3D wide baseline), whose views correspond to the right-most and the left-most views of the recommended 3 views in [15].

In 3D-HEVC compression scheme, the inter-view correlation between two views is explored to improve compression efficiency. Specifically, one view is encoded as base view using HEVC standard, and the dependent view is encoded by exploring inter-view prediction and the inter-frame prediction as well. To generate asymmetric compression artifacts, we set the quantization parameter (QP) of base view and dependent view differently. We further put compression artifacts against different depth levels to discover the visibility of 3D-HEVC artifacts in different presence of depth (i.e., different baseline groups in TABLE I) . To create the PVSs with zero baseline, the same camera video is used to create the distorted videos of left view and right view. The video is essentially compressed with HEVC to generate different artifact levels.

C. Subjective Test Methods

Depth sensation and the annoyance of compression distortion are two major factors that influence viewer's experience of quality, when one watches 3D-HEVC compressed video. The video quality and depth quality were measured using single stimulus method ACR-HR on 5 discrete scales with 1 for bad quality and 5 for excellent quality, according to the recommendation in ITU-R BT.2021 [4]. The viewers were instructed that video quality refers the perceived quality of the pictures, and depth quality refers to the ability of the video to deliver an enhanced sensation of depth. Moreover, we measured the overall quality of experience using ACR-HR, and viewers were instructed in the training phase to comprehensively consider the video quality and depth quality, but not limited to these.

28 non-expert student subjects (assessors) aged from 21 to 26 years old participated in the subjective test. All participants had a visual acuity above 1.0 and passed the color vision test (Ishihara plats) and stereo vision test (RANDOT). The 3D display is 55-inches SAMSUNG UA55HU8500J 3D television with shuttle glasses. The viewing distance is about 2 meters. Before the formal test, there are an instruction and a training phase to make a subject get familiar with the test and establish a stable assessment criteria. The formal subjective test including 156 PVSs is split into three sessions to avoid the viewers being fatigue or bored. The display order of PVSs was random for each subject. When a PVS was displayed, there was an interface for the subject to input his assessment of video quality, depth quality, and overall QoE sequentially. The rating interface was implemented properly so that no viewing discomfort or mode-switch interruption were introduced.

Post-processing of the collected quality scores was performed to screen out the subject whose correlation with the average scores of all subjects was lower than 0.75. Finally, there are 22 valid subjects (13 males and 9 females).

III. Analysis of 3D-HEVC Database

Firstly, we analyze the suitability of the evaluation methods used in creating the database. Secondly, we investigate the influential factors to the multiple aspects of QoE for 3D videos, and draw some conclusions based on the observations on our database.

A. Quality of Original Sequences

In Fig. 3, for each original video content with different camera baselines (i.e. HRC001, HRC101, HRC201), the video quality, depth quality, and overall quality are shown separately in sub-figure (a)(b)(c). All original videos have good video quality (MOS $>$ 3.5) as source videos for creating database, as shown in Fig. 3(a). Controlling camera baseline indeed effectively changes depth sensation, as shown in Fig. 3(b). Moreover, it can be observed that 2D videos without disparity also deliver some depth sensation depending on the characteristics of the video content. Detail analysis will be done in subsection E.

It is interesting to notice that video quality varies little with the increase of depth sensation, as shown in Fig. 3(a). However, depth sensation indeed contributes positively to the overall quality of experience as shown in Fig. 3(c), especially when comparing 3D video with 2D video. Disparity information produces a compelling sense of depth, which defines the added value of stereoscopy.

B. Subjects' agreement

For each PVS, 22 valid subjects gave their opinion scores. We calculate the 95% confidence interval (CI95) according to

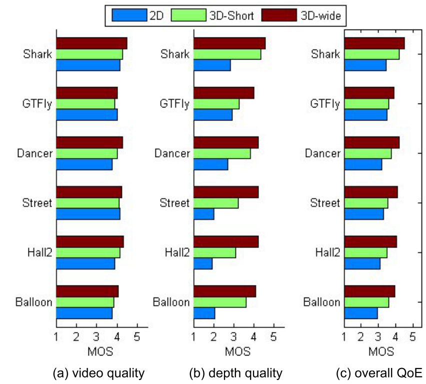

Fig. 3. MOS for original monoscopic and stereoscopic sequences

TABLE II. Average of 95% CI of 156 PVSs

Video quality	Depth quality	Overall quality
0.31	0.37	0.30

ITU-T BT500, which is influenced by the opinion variation between subjects. It means that the difference between the experimental mean score and the 'true' mean score (for a very high number of subjects) is smaller than the CI95 with a probability of 95%. The average of CI95 for all PVSs is shown in Table II separately for the video quality, depth quality, and overall QoE. It can be seen that the subjects have arrived at a reasonable agreement on the perceptual quality. Therefore, the MOS values obtained can be regarded as the ground truth.

C. Relationship of Multidimentional Qualities

In Fig. 4, we average the MOS across six video contents for each HRC. The upper bound of the interval around an average MOS point corresponds to the maximum MOS of the six video contents, and the lower bound corresponds to the minimum.

We can observe that, first, video quality is influenced more by compression levels and video content than by the varied depth ranges (i.e. zero baseline, short baseline, wide baseline), as shown in Fig. 4(a). Second, depth quality varies only slightly with video compression levels with a maximum

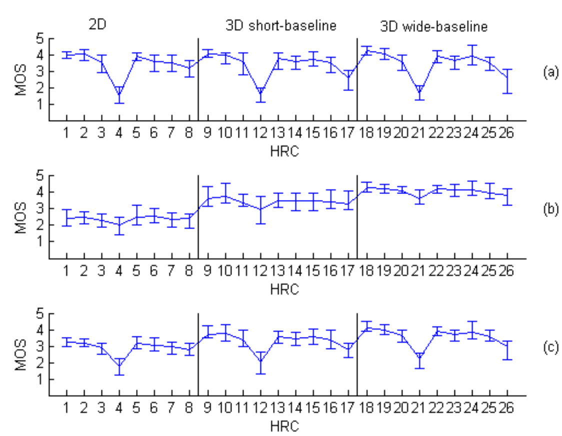

Fig. 4. (a) MOS of video quality for each HRC; (b) MOS of depth quality for each HRC; (c) MOS of overall quality for each HRC

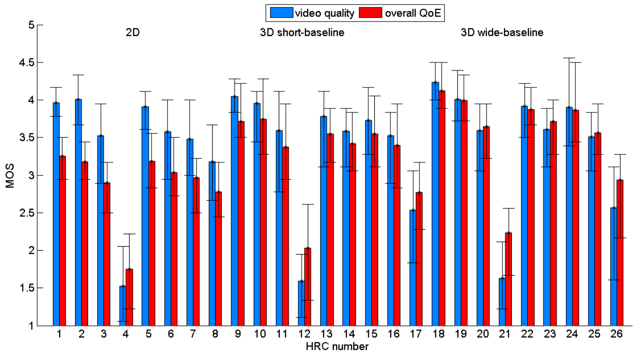

Fig. 5. Compare video quality with overall quality

TABLE III. CORRELATION BETWEEN THE QUALITY DIMENSIONS

Video quality vs. depth quality	Video quality vs. overall QoE	Depth quality vs. overall QoE
0.2587	0.8936	0.6412

MOS difference less than 1 point, given a camera baseline, as shown in Fig. 4(b). Third, overall quality in Fig. 4(c) is jointly determined by video quality and depth quality, but compression artifact is a dominant factor to the QoE compared to the varying depth range. For 156 PVSs, we calculate the correlation between video quality, depth quality, and overall quality, which is shown in Table III. The results confirm these observations.

It is worth noting that the added value of depth sensation also exists even in 3D videos with poor compression artifacts, e.g. HRC number 12 and 21, as can be observed by comparing Fig.4(a) and Fig.4(c). This observation contradicts the conclusion in [17] that the experienced added value of stereoscopic depth is visible only if the artifact level is low.

To investigate this discrepancy, we draw Fig. 5 to better illustrate the rating difference between video quality and overall QoE. Some interesting conclusions can be drawn. Firstly, subjects assessing overall QoE tend to be reluctant to give higher score than video quality score, when the video quality is fairly good (e.g. above 3 points). A potential reason is that subjects tend to reserve some margin for considering depth quality when they use a comprehensive criteria to assess overall QoE. This explanation can be further confirmed by the observation that the overall quality score is increased to emulate the video quality when depth range is increased from zero baseline to wide baselines. Secondly, subjects tend to be reluctant to give worse overall QoE score when the video quality is poor (below 2.5 points), for the same reason of comprehensive criteria, by considering the added value of depth sensation. This explanation is similarly confirmed by the fact that the difference between overall QoE and video quality increases with the increased baselines when the video quality is poor, as shown by HRC number 4, 12, 21.

Notice that a novel quality evaluation method was used in [17] to combine a conventional quantitative psychoperceptual evaluation and a qualitative descriptive quality evaluation based on the individual's own vocabulary. Therefore, the discrepancy may arise from different evaluation methods and different interpretation of quality scales.

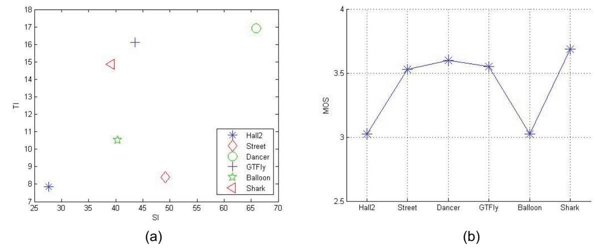

Fig. 6. The influence of video characteristics on video quality at large, (a) SI/TI of video contents, (b) variation of MOS with video contents

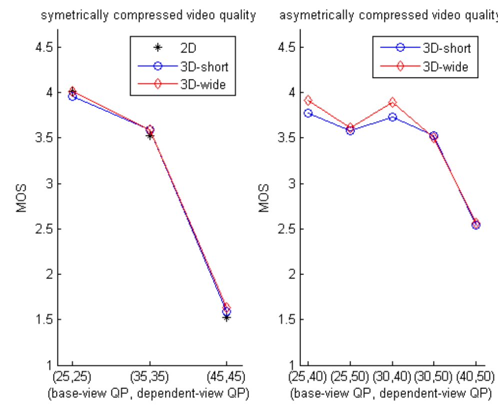

Fig. 7. Video quality of symmetrically and asymmetrically compressed video at different depth range

D. Key Factors to Video Quality

Video quality is dominantly determined by the 3D-HEVC compression levels as shown in Fig. 4(a). It is secondarily influenced by the characteristics of video content. The MOS difference between different video content under the same QP condition can be as large as up to 1.5 points as shown, for examples, by HRC number 11, 17, 26 of Fig. 4(a).

To understand how video characteristics may influence video quality, we average the MOSs of PVSs of each video content across the HRCs, and derived Fig. 6(b). For the convenience of readers, Fig. 1(a) is repeated in Fig. 6(a). It can be seen that video contents having either larger SI or larger TI values (e.g. 'Shark', 'Dancer','GTFly') tend to have higher video quality under the same QP settings. Video content "Hall" having small SI and TI value has lower quality. This may be due to the fact that the compression artifact in the smooth area of the video is more visible under the same QP settings.

Video quality is hardly influenced by the varied depth range in the symmetric compression case, no matter the QP is small or large, as shown in Fig. 7. However, in the asymmetric compression case, it is slightly better in the wider depth range case when the QPs of both view are not larger than 40.

E. Key Factors to Depth Sensation

Since depth quality varies only slightly with HRCs having different compression levels as shown in Fig. 3(b), we derived the average MOS of each video content across HRCs and obtained Fig. 8. It is clear that the characteristics of video content and depth range are two main factors to depth quality variation.

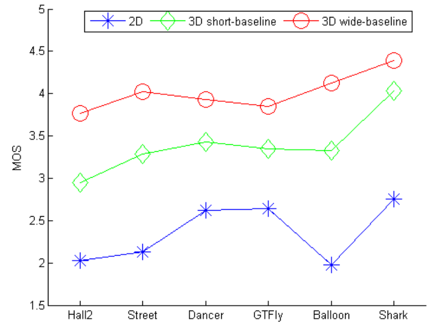

Fig. 8. Depth quality varies significantly with video contents and the camera baselines

Disparity significantly increases the depth sensation as shown by comparing the green curve with the blue curve. Besides, the 2D video without disparity also presents depth cues. For example, the motion information may lead to the sense of perspective, which depends heavily on content characteristics.

IV. Proposed Bitstream-level Quality Measure

A. Proposed Metric

Based on our 3D-HEVC video database, we propose a no-reference 3D-HEVC bitstream-level quality assessment model to assess the perceived quality of the stereoscopic video. The block diagram of the proposed model is shown in Fig. 9. It consists of three steps.

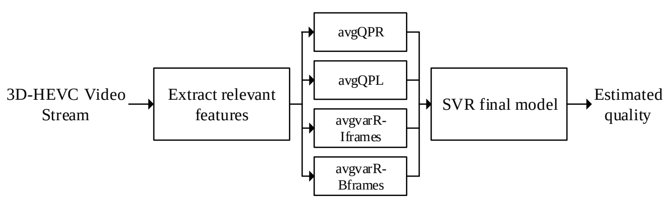

Fig. 9. The block diagram of the proposed NR VQA model

Firstly, we parse the 3D-HEVC video bit-stream and extract some parameters such as quantization parameters (QP) of the base view and the dependent view, and the variance of the prediction residuals of coding blocks (CU) of base view in the luminance channel. Secondly, four features are derived from the extracted parameters, namely, the average QP of base view (avgQPR), the average QP of dependent view (avgQPL), average variance of CUs' prediction residuals of I-frames (avgvarR-Ifrms), and the average variance of CUs' prediction residuals of B-frames at the second inter-frame prediction level of the GOP structure (avgvarR-Bfrms). Finally, we put all features into the Support Vector Regression (SVR) model to predict the ultimate objective video quality:

$$Quality = SVR(QPR, QPL, varRIframes, varRBframes) \qquad (1)$$

SVR maps features into higher-dimensional space and predicts the 3D video quality [18]. We use the ϵ-SVR with the radial basis kernel function and the loss function parameter to 0.25.

According to the subjective study of 3D video quality in Section III, the dominant factor to 3D video quality is compression levels, which is essentially controlled by QP settings of base view and dependent view. The secondary factor is the characteristics of video content. We find that the average variance of I-frames of base view reflect the texture complexity of video content to a degree. Whereas, the average variance of B-frames of base view reflect the temporal complexity of video content to some degree. With these four key features, our preliminary model is expected to deliver reasonably good performance.

B. Performance Analysis

We did 100 times of cross validation with 80% samples randomly selected for training and other 20% for test on the 3D-HEVC database. The performance is evaluated by computing the Spearman Rank Order Correlation Coefficient (SROCC) [15], the Pearson Correlation Coefficient (PCC), and the Root Mean Square Error (RMSE) between predicted score and MOS. SROCC assesses how well the relationship between two variables can be described using a monotonic function. PCC measures the linear relationship between a models performance and the subjective data. RMSE provides a measure of the prediction accuracy. SROCC and PCC takes value from $[-1, 1]$, with the values close to 0 declaring very bad or no correlation and the values close to 1 representing high positive correlation.

The performance of the proposed model is shown in Table IV. The median, mean, and standard deviation of the SROCC, PCC, and RMSE values of 100 times cross validation are reported. It can be seen that the model achieves a fairly good performance on our database.

TABLE IV. Performance of the proposed NR method

	Median value	Mean value	Standard deviation
SROCC	0.864	0.846	0.079
PCC	0.922	0.914	0.044
RMSE	0.400	0.407	0.103

C. Comparison with FR metrics

In general, a good Full-Reference perceptual quality assessment metric should be more accurate than a No-Reference method. Hence, we compared the proposed bitstream-layer No-Reference model with some widely used Full-Reference 2D quality assessment metrics, such as PSNR, FSIM [19], MS-SSIM [20] and VQM (i.e. ITU-T Rec. J.144) [21]. Since PSNR, SSIM, FSIM, MS-SSIM are image quality metrics, SROCC of a 3D video is obtained by averaging SROCCs across all frames of a PVS. VQM is a 2D video quality assessment method, thus, we averaged SROCCs of left-view and right-view videos to derive the SROCC of a 3D video. As can be seen in Table V, on our 3D-HEVC video database, the proposed bitstream-layer No-Reference model performs better than these classic 2D Full-Reference perceptual video quality assessment metrics.

Further, we compared the performance of the proposed model with a state-of-the-art pixel-level Full-Reference 3D video assessment method [9]. The key idea of the method in [9] is a 2D-to-3D weighting scheme that accounts effectively binocular perception mechanism of human visual system. The SROCC performance by applying the weighting scheme to the widely-used FR 2D metrics is given also in Table V. All of

TABLE V. SROCC PERFORMANCE OF PIXEL-LEVEL FR METRICS ON OUR 3D-HEVC DATABASE

Method	Direct Average	Weighted Average[9]
PSNR	0.482	0.483
SSIM	0.415	0.416
FSIM	0.814	0.816
MS-SSIM	0.689	0.690
VQM	0.771	0.772
Proposal	**0.846**	

the frame-based video quality assessment method are using the energy-based weighting algorithm on each frame except for VQM, which produces a sequence-level quality score and use the weighting algorithm at sequence level.

It can be seen that the performance of the proposed bitstream-layer 3D video quality assessment model is more effective than the state-of-the-art Full-Reference video quality assessment based on the pixel domain. This result is reasonable because the key features for compression artifacts are QP and video complexity, which can be more easily and accurately obtained at the bitstream level than at the media level (i.e. in the pixel signal domain). We made the performance comparison only on our database, because it is currently the only 3D-HEVC video database with compressed bitstream so far as we know.

V. CONCLUSION AND FUTURE WORK

We build a 3D-HEVC video database and propose a No-Reference bitstream-level objective quality assessment model based on an analysis of the key factors to video quality. The experiment results validate the effectiveness of the model by comparing it with some state-of-the-art 2D and 3D Full-Reference pixel-level objective quality assessment metrics.

In the future, we will further explore features from inter-frame and inter-view prediction to improve our model. More state-of-the-art 2D and 3D metrics will be taken into consideration for the comparative study of model performance.

Moreover, we will extend our database to investigate the influence of 3D-HEVC compression on the synthesized view and make the whole 3D-HEVC bitstream video database public to the research community. We will also consider using the Pair Comparison method [22] for the assessment of QoE, which has the advantage of avoiding multiple quality dimensions and interpretation issues.

ACKNOWLEDGMENT

This work was supported by the NSFC (Natural Science Foundation of China) under Grant 61571413, and National 973 Program of China under Grant 2015CB351803, and in part by NSFC Grant 61390514 and Intel ICRI MNC.

REFERENCES

[1] G. Tech, Y. Chen, K. Muller, J.-R. Ohm, A. Vetro, and Y.-K. Wang, "Overview of the multiview and 3d extensions of high efficiency video coding," 2015.

[2] M. H. Pinson, N. Staelens, and A. Webster, "The history of video quality model validation," in *Multimedia Signal Processing (MMSP), 2013 IEEE 15th International Workshop on*. IEEE, 2013, pp. 458–463.

[3] C.-C. Su, A. K. Moorthy, and A. C. Bovik, "Visual quality assessment of stereoscopic image and video: challenges, advances, and future trends," in *Visual Signal Quality Assessment*. Springer, 2015, pp. 185–212.

[4] I. Recommendation, "2021, subjective methods for the assessment stereoscopic 3dtv systems," *Internationale Telekommunikations-Union, Geneva, Switzerland*, 2012.

[5] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Quality assessment of stereoscopic images," *EURASIP journal on image and video processing*, 2008.

[6] M.-J. Chen, L. K. Cormack, and A. C. Bovik, "No-reference quality assessment of natural stereopairs," *Image Processing, IEEE Transactions on*, vol. 22, no. 9, pp. 3379–3391, 2013.

[7] X. Wang, M. Yu, Y. Yang, and G. Jiang, "Research on subjective stereoscopic image quality assessment," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2009, pp. 725 509–725 509.

[8] J. Wang and Z. Wang, "Perceptual quality of asymmetrically distorted stereoscopic images: the role of image distortion types," in *Proc. Int. Workshop Video Process. Quality Metrics Consum. Electron*, 2014, pp. 1–6.

[9] J. Wang, S. Wang, and Z. Wang, "Quality prediction of asymmetrically compressed stereoscopic videos," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 3427–3431.

[10] P. Seuntiens, L. Meesters, and W. Ijsselsteijn, "Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric jpeg coding and camera separation," *ACM Transactions on Applied Perception (TAP)*, vol. 3, no. 2, pp. 95–109, 2006.

[11] S. Jumisko-Pyykkö, T. Haustola, A. Boev, and A. Gotchev, "Subjective evaluation of mobile 3d video content: depth range versus compression artifacts," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2011, pp. 78 810C–78 810C.

[12] V. De Silva, H. K. Arachchi, E. Ekmekcioglu, and A. Kondoz, "Toward an impairment metric for stereoscopic video: A full-reference video quality metric to assess compressed stereoscopic video," *Image Processing, IEEE Transactions on*, vol. 22, no. 9, pp. 3392–3404, 2013.

[13] M. Urvoy, M. Barkowsky, R. Cousseau, Y. Koudota, V. Ricorde, P. Le Callet, J. Gutiérrez, and N. Garcia, "Nama3ds1-cospad1: Subjective video quality assessment database on coding conditions introducing freely available high quality 3d stereoscopic sequences," in *Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on*. IEEE, 2012, pp. 109–114.

[14] L. Goldmann, F. De Simone, and T. Ebrahimi, "A comprehensive database and subjective evaluation methodology for quality of experience in stereoscopic video," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2010, pp. 75 260S–75 260S.

[15] D. Rusanovskyy, K. Müller, and A. Vetro, "Common test conditions of 3dv core experiments," *ITU-T SG*, vol. 16, 2013.

[16] "Final report from the video quality experts group on the validation of objective models of multimedia quality assessment, phase i [online]," *Available: ftp://vqeg.its.bldrdoc.gov/Documents/*.

[17] D. Strohmeier, S. Jumisko-Pyykkö, and K. Kunze, "Open profiling of quality: a mixed method approach to understanding multimodal quality perception," *Advances in multimedia*, vol. 2010, 2010.

[18] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004.

[19] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: a feature similarity index for image quality assessment," *Image Processing, IEEE Transactions on*, vol. 20, no. 8, pp. 2378–2386, 2011.

[20] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*, vol. 2. Ieee, 2003, pp. 1398–1402.

[21] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *Broadcasting, IEEE Transactions on*, vol. 50, no. 3, pp. 312–322, 2004.

[22] J. Li, M. Barkowsky, and P. Le Callet, "Validation of reliable 3dtv subjective assessment methodology-establishing a ground truth database," *VQEG eLetter*, vol. 1, no. 2, pp. 30–35, 2014.