INTERNATIONAL SCIENTIFIC AND VOCATIONAL JOURNAL (ISVOS JOURNAL)

Vol.:2 Issue:2 Date: December 2018 Received: 08.12.2018 Accepted: 14.12.2018 Final Version: 27.12.2018

ISVOS Journal 2018, 2(2):99-103

A Reduced Reference Metric for Enhanced 3D Video Perception

Gökçe Nur Yılmaz[a,1]

[a]Electrical and Electronics Engineering Department, Kirikkale University, Kirikkale, Turkey

---

Abstract

Currently, one of the trending research topics among the researchers assisting to the enhancement of the 3D video services relies on the 3 Dimensional (3D) video Quality of Experience (QoE) prediction metric development process. The researches for this trending topic can be ensured by characterizing the 3D video related features in the most compatible way as possible in this process. Considering this fact, a novel Reduced Reference (RR)) 3D video QoE prediction metric relying on color+depth map 3D video representation is developed in this study. The developed metric utilizes the incorporation of the Significant Information (SI) in the depth map videos with the Structual Complexity Information (SCI) of their color counterparts. The abstraction filter and Structural SIMilarity Index (SSIM) are exploited for the SI and SCI computations, respectively. Performed subjective experiment results are utilized to predict the performance of the developed metric. Observing highly effective results after the performance evaluation process, it can be clearly stated that the developed RR metric is compatible for assisting the advancement of the 3D video services.

Key Words: "3D video, depth level, structural complexity"

---

1. Introduction

Developments in 3 Dimensional (3D) video technologies lead 3D video services to enable better Quality of Experience (QoE) to the end users. Improving this QoE can only be supported by exploiting metrics including 3D video characteristics associated with Human Visual System (HVS) features. The predicted 3D video QoE can then be enhanced by the 3D service providers [1].

Currently, this support is provided by subjective tests considering real observers during the 3D video QoE assessment. However, these tests are cost and time ineffective. Therefore, mathematical model based objective metrics should be used since these metrics are quick and iterative during the QoE assessment process. Therefore, the 3D service providers are in high demand for the objective metrics reliable and efficient to predict the 3D video QoE [2].

The objective QoE metrics are split into 3 types in the literature. The names of these types are Full Reference (FR), Reduced Reference (RR) and No Reference (NR). As the names indicate, the FR metric type requires full information regarding the original and encoded videos for the QoE evaluation. The RR type, on the other hand, only need extra information derived from the original and/or encoded videos. However, the NR metric type does not require any information associated with the original video for the QoE evaluation. Envisaging that the RR metric type includes only important information derived from the original and/or encoded videos for the QoE evaluation [3], it is used as metric type in this study.

There exist different 3D video representation forms in the literature. Considering the advantages of the color+depth map 3D video representation [3], the metric is developed using this form.

Considering these facts, a RR objective 3D video QoE metric using the association of the Significant Information (SI) of the depth maps with the Structural Complexity Information (SCI) of their color video counterparts is proposed in this study. The remarkable depth levels used to stress the objects and background are indicated by the SI for the depth map video contents. The SI of the original and compressed depth map videos are obtained by abstracting them using the bilateral filter [4][5]. The SI results are exploited as extra information in the proposed RR metric. For the SCI measurement, Structural SIMilarity Index (SSIM) metric [6] is extended to measure the differences between the original and encoded color videos. The SCI measurement results are also used as extra information in the developed RR metric.

The subjective experiment results obtained for the performance prediction of the developed metric indicate the compatibility of it to predict the QoE of the 3D videos.

[1] Corresponding author. Tel.: 0090 318 357 42 42-1207; fax: +0-000-000-0000 .
E-mail address: gnur@kku.edu.tr

The remaining parts of the paper is organized as following. Section 2 discusses the RR 3D video QoE metrics existing in the literature. Section 3 represents the proposed metric. The performance evaluation results are shown in Section 4. Finally, the paper concslusion is given in Section 5.

2. Related Work

In this section, the existing RR metrics in the literature are presented. In [6], edge-information, which represents the basic structure of the depth map contents, is used as extra information to propose a global RR objective metric for the color plus depth map 3D video quality evaluation as an extension of the studies in [7] and [8]. Moreover, a RR 3D video perception metric proposed based on structural distortions measured using edges and contour information difference between original and compressed depth map sequences is discussed in [9] as an extension of the study in [6]. A perceptual RR metric is proposed based on comparing the edge information in the distorted video with the original one in [10]. In [11], the authors utilizes pixel neighbourhood and edge information of the objects associated with the color+depth map 3D video. It should be stressed here that the incorporation of the structural information with the SDL is not used before for the 3D video QoE prediction in the literature.

3. Proposed Metric

The framework of the proposed metric is shown in Figure 1. In order to develop the metric, first of all, the color and depth map sequences of the ten 3D video contents are separately encoded. The names of these contents are Advertisement, Breakdance, Butterfly, Windmill, Chess, Interview, Farm, Football, Newspaper, and Ballet. Throughout the encoding process, JSVM 9.13.1 codec [13] is utilized. 5 different QPs ranging from 25 to 45 with a sequential order of 5 are used to encode the contents. Moreover, 25 fps and 1920x1080spatial resolutions are used as the encoding parameters. Considering the 5 QPs used to encode each of the color and depth map videos belonging to 10 different 3D video contents, 250 encoded 3D videos are obtained from their combinations to develop the proposed metric. Then, the SI of these 250 compressed videos and 10 original associated versions of them are separately computed by utilizing the abstraction filter [4]. The bilateral filter [5] relying on Gauss filtering method is used as the abstraction filter since it is quite effective for the depth map videos. Using this filter, the significant/meaningful information for the HVS is stressed in these contents. This stressed information is considered as the HVS related side information that is exploited in the RR metric. The SI for the original and compressed depth map video contents indicates the remarkable depth levels used to stress the objects and background.

In parallel with the SI calculation as the extra information, the SCI representing the structural differences between the original and encoded color video contents are computed using the SSIM. This difference measurement is again exploited as HVS related extra information in the developed metric. The reason of utilizing the SSIM for the SCI measurement relies on the fact that it is widely utilized metric among the researchers and it is quite efficient to measure the structural information. As can be observed from Figure 1, the SI measurement between the abstracted original and encoded/compressed depth map videos is named as RDL_D in the developed metric. Moreover, the SSIM measurement relying on the SCI differences between the original and compressed color contents are named as $SSIM_C$. By combining RDL_D with $SSIM_C$ as follows, the developed metric is obtained.

$$3DM_{CD} = \lambda RDL_D SSIM_C \tag{1}$$

where, $3DM_{CD}$ presents the proposed metric. In order to keep the results of the proposed metric between 0 and 5, the λ constant is added to it. According to the analysis results of the 3DMCD results obtained from the 250 3D video contents and their associated 10 original ones as 2.23x102.

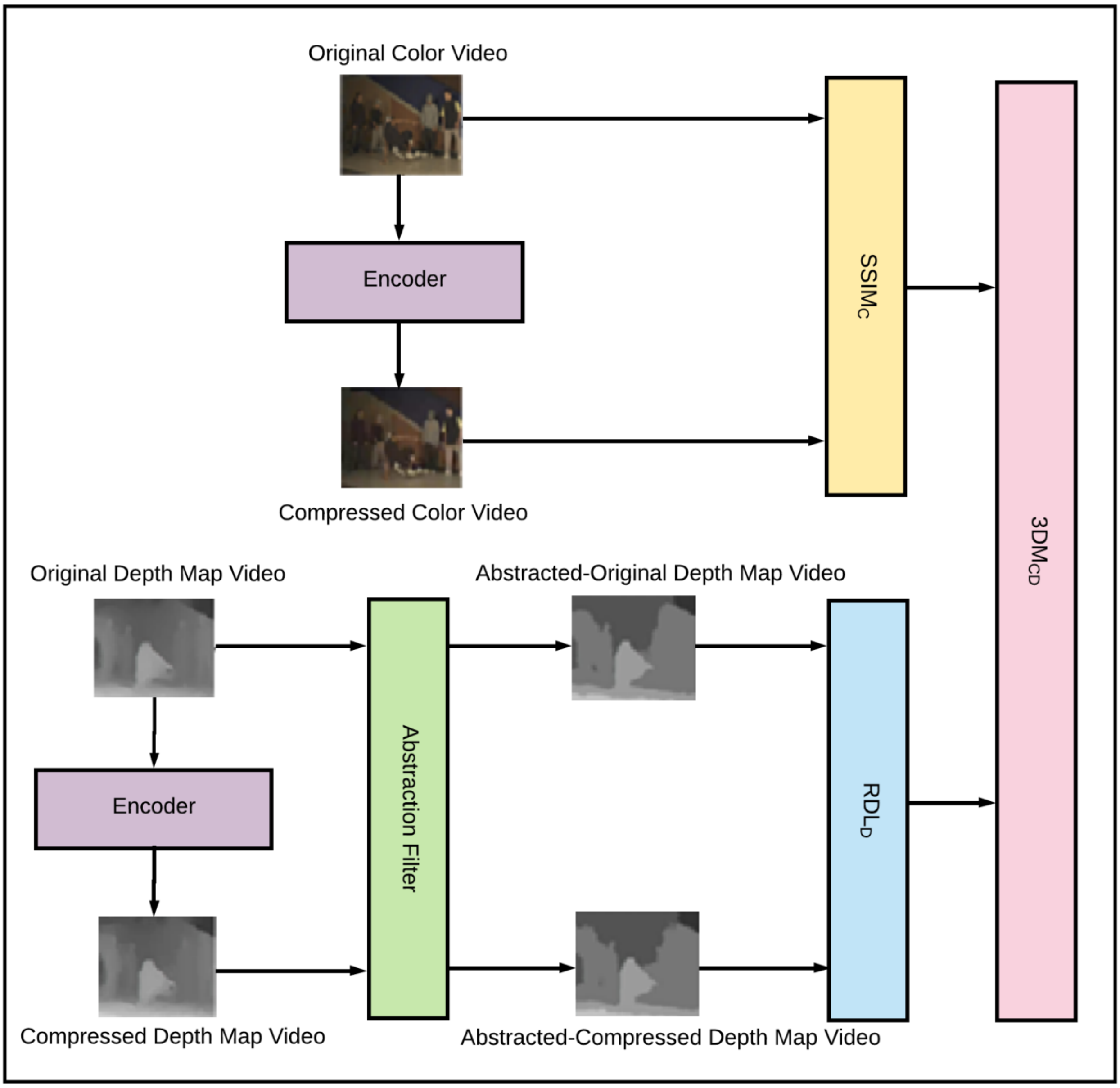

Figure 1: The diagram of the developed metric

4. Results and Discussions

During the performance evaluation process, 5 3D video contents encoded to develop the proposed metric and their original ones are used (i.e., Advertisement, Breakdance, Butterfly, Windmill, Chess). Therefore, 125 encoded 3D video contents are used to evaluate the proposed metric. Then, the $3DM_{CD}$ results for each of these 125 contents are computed. In order to prove the compatibility of the developed metric, subjective experiments are also conducted with these 125 encoded 3D videos. A ranking scale ranging from 1 to 5 are utilized throughout the experiments. 1 and 5 represent the lowest and highest QoE results, respectively. An autostereoscopic display is exploited to show the 3D videos to the viewers. The display is of 39 inch. SSCQE [14] is considered as the subjective test method for the experiments. The viewers are asked to evaluate the overall 3D video satisfaction including both the 3D video quality and depth perception during the experiments. 23 viewers joined to the tests. The ages of these 23 viewers range from 18 to 31. After determining 2 viewers as outliers, the Mean Opinion Scores (MOSs) of 21 viewers are computed as advised by the ITU-R BT 500-13 [14]. Then, the relationship between the MOSs and $3DM_{CD}$ results are correlated by Symmetrical Logistic Function (SLF) [14]. Then, Correlation Coefficient (CC) results are computed using the SFL results. The CC results range from 0 to 1. A CC result closer to 0 represents low correlation and vice versa. The averages of the CC results are then calculated for each of the 5 3D video contents (i.e., Chess, Advertisement, Ballet, Interview, Breakdance, and

Farm). Moreover, Root Mean Square Error (RMSE) results are computed using the results of the SFL. In contrast to the CC results, a RMSE result closer to 1 leads to high errors and vice versa.

Table 1 presents the average CC and RMSE results of the 5 encoded 3D videos. As can be seen from the results in the table, high correlation (close to 1) is observed for all of the contents. Moreover, low errors (close to 0) are obtained for the contents. For instance, as seen from the table, the average CC results for the Ballet 3D video content is 0.974. The average RMSE results for the same content is 0.0281.Similar patterns are observed for the rest of the 3D video contents. These results present the compatibility of the developed metric for QoE assessment of the 3D videos.

Table. 1. The performance evaluation results of the developed metric

3D Video Content	CC	RMSE
Advertisement	0.982	0.0252
Ballet	0.974	0.0281
Interview	0.961	0.0354
Breakdance	0.967	0.0313
Farm	0.978	0.0267

5. Conclusion

Developed is a RR 3D video QoE evaluation metric. The remarkable depth level features and structural information important for the HVS have been accompanied in the developed metric. Good correlation results obtained between the conducted subjective experiment and developed metric results have proven the metric's performance compatibility. In the future studies, more HVS related features will be incorporated with the developed metric to boost its effectiveness.

Acknowledgement

This work has been supported by the Scientific and Technological Research Council of Turkey, Project Number: 114E551.

References

[1] Q. Huynh-Thu, P. Le Callet, and M. Barkowsky, "Video Quality Assessment: From 2D to 3D Challenges and Future Trends," 17th IEEE International Conference on Image Processing, Hong Kong, 26-29 Sep. 2010.

[2] Chaminda T. E. R. Hewage, "Perceptual Quality Driven 3D Video over Networks," PhD Thesis, Centre for Communication Systems Research Faculty of Engineering and Physical Science, University of Surrey, UK, 2008.

[3] C.T.E.R. Hewage, S. T.Worrall, S.Dogan, S. Villette, A. M. Kondoz, "Quality Evaluation of Color Plus Depth May Based 3D Video," IEEE Journal of Selected Topics in Signal Processing, vol. 3., no. 2, Apr. 2009.

[4] G. Nur Yilmaz, "A Depth Perception Evaluation Metric for Immersive 3D Video Services," 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video, 7-9 June 2017.

[5] L. Caraffa, J. P. Tarel, P. Charbonnier, "The Guided Bilateral Filter: When the Joint/Cross Bilateral Filter Becomes Robust," IEEE Transactions on Image Processing, 24 (4), pp 1199-1208, 2015.

[6] Chaminda T.E.R. and Martini M. G., "Reduced-Reference Quality Assessment for 3D Video Compression and Transmission," IEEE Transactions on Consumer Electronics, vol. 57, issue 3, pp. 1185-1193, Aug. 2011.

[7] C. Hewage and M.G. Martini, "Reduced-Reference Quality Metric for 3D Depth Map Transmission," 3DTV Conference, Tampere, Finland, June 2010.

[8] C. Hewage and M.G. Martini, "Reduced reference image quality metric for compressed depth map associated with colour plus depth 3D video," IEEE International Conference on Image Processing (ICIP), Hong Kong, 26-29 Septermber 2010.

[9] Chaminda T.E.R., Martini M. G. 2012 "Edge-based Reduced-Reference Quality Metric for 3-D Video Compression and Transmission," IEEE Journal of Selected Topics in Signal Processing, 5, 471-482.

[10] Chaminda T.E.R. and Martini M. G., "Edge-based Reduced-Reference Quality Metric for 3-D Video Compression and Transmission," IEEE Journal of Selected Topics in Signal Processing, vol. 6, no.5, Sep. 2012.

[11] Martini M. G., Chaminda T.E.R., and Villarini B., "Image Quality Assessment Based on Edge-Preservation," Signal Processing: Image Communication, vol. 27, issue 8, pp. 875-882, Sep. 2012.

[12] C. Fehn, "Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV", Proc. of SPIE Conf. Stereoscopic Displays and Virtual Reality Systems XI, vol. 5291, pp. 93 – 104, CA, U.S.A., January 2004.

[13] JSVM 9.13.1. CVS Server [Online]. Available Telnet: garcon.ient.rwth aachen.de:/cvs/jvt

[14]https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-13 201201-I!!PDF-E.pdf