No Reference Objective Quality Metric for Stereoscopic 3D Video

Yi Han[1,2], Zhenhui Yuan[2], Gabriel-Miro Muntean[2]
[1] School of Computer Science, University College Dublin, Dublin, Ireland
[2] School of Electronic Engineering, Dublin City University, Dublin, Ireland
yi.han@ucdconnect.ie, {zhenhui.yuan, gabriel.muntean}@dcu.ie

***Abstract*— The stereoscopic three-dimensional (3D) video technologies have achieved significant success in providing enhanced immersive experience to consumers. However network delivery of 3D video content at good quality levels is challenging mostly due to the variable network conditions. In this context, efficient objective 3D video quality assessment is a critical aspect, in particular for video service providers who need to adjust the video delivery process to the network conditions in real-time. Current objective 3D video quality assessment methods are reference-based, requiring the availability of the original 3D video sequences, which is difficult to achieve in practice. Additionally, most of the existing 3D video quality metrics are developed for depth-enhanced 3D. This paper proposes the No reference objective Video Quality Metric (NVQM) for real-time 3D video quality assessment. NVQM considers the correlation between network packet loss and perceptual video quality for different bit-rate video sequences. NVQM is modeled based on the video quality model specified in ITU-T G.1070 and tuned according to results of extensive subjective tests. NVQM was developed for the evaluation of side-by-side stereoscopic 3D sequences, the most widely commercialized 3D video format. The performance of NVQM is studied by comparing against three state-of-the-art video quality objective models: structural similarity index (SSIM), video quality metric (VQM), and ITU-T G.1070. Results show that NVQM outperforms the existing objective metrics with up to 23% in terms of accuracy.**

Keywords—3D video; objective quality assessment; non-intrusive; stereoscopic

I. Introduction

Currently, the three-dimensional (3D) video is gaining increasing popularity by providing immersive user experience. Enhanced from the conventional 2D format, the 3D video excels at bringing an almost live scene closer to the users, introducing them into the original environment of the displayed content. With the help of image processing and advanced filming and display technologies, 3D movies have drawn very much attention from the audience and made significant profits in cinemas worldwide. Similarly, 3D content has also attracted great interests in other application areas, such as 3DTV [1], 3D gaming, 3D conferencing, etc.

Thanks to the rapid development of digital video compression and transmission technologies which includes H.264/AVC, H.264/SVC and multiview video coding (MVC) [2], new real-time 3D video applications appear on the market with enhanced interactive capabilities. The perceptual 3D video quality is a major factor in assessing these applications. Limited research has been conducted to measure the quality of experience (QoE) levels of the 3D video. Employing subjective methods for evaluating 3D video quality provides the most accurate results as it reflect directly human perception levels. However, they are time consuming and human resource intensive. Additionally, subjective quality assessments require controlled environments, unavailable in real-time remote delivery of 3D video content, so objective metrics are preferred to be used. Recently, several objective assessment methods have been proposed in [3] - [8], but they lack the accuracy of assessing stereoscopic 3D video. This is due to the fact that the human visual system (HVS) is difficult to model using pixels and depth, and is also affected by human eye comfort level, viewing distance, etc. Furthermore, the existing objective 3D video quality assessment methods are highly dependent on the original video content and none of them directly considers network impairments.

Video encoding and decoding process causes quality degradation and the transmission process affects additionally the content quality, due to network delivery effects such as packet loss, delay and jitter. There are several widely employed 2D video quality metrics such as PSNR [9], SSIM [10], VQM [11], no reference PSNR [12] which can be used for assessing 3D video quality approximately [3] [4]. However, current objective assessment methods are intrusive [13], which means that they require the usage of the original 3D video content in the full reference methods. These objective methods need to analyse the decoded video content in order to assess their blockiness, blurring, and depth information. Such assessment can only be done off-line rather than in "real-time" during the video transmission.

Figure 1 shows a typical scenario of delivering 3D video content. The video quality degradation is caused by both codec and transmission processes. In this context, the relationship between the network conditions and the resulting 3D video quality is required to be studied. Additionally, it has been proved that both bit-rate and frame-rate have significant impact on the 3D video quality [3]. In this paper we propose a novel No reference Video Quality Metric for 3D video quality assessment (NVQM) that considers both bit-rate and network characteristics (i.e. packet loss) as input. Based on a new model, NVQM correlates network

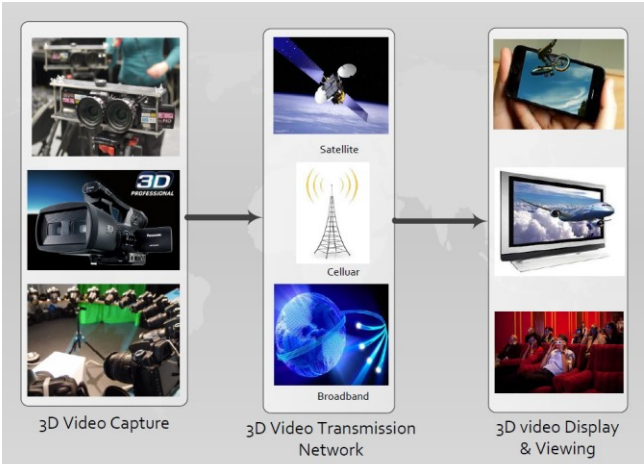

Figure 1. Phases in 3D Video Transmission

conditions to the 3D video quality and relies on video encoding settings only, without the need of processing in details the video images. The advantage of NVQM model is that it can be used as an estimation of the 3D video quality in real-time during the network transmission. The model is of great benefit to the 3D community, especially for adaptive 3D video transmissions, for instance, as the video quality at the client can be predicted based on the delivery conditions and the service provider can then pro-actively decide to adjust the bit-rate and/or network bandwidth, rather than reactively, being more efficient. This cannot be done using any current objective 3D quality assessment methods.

The remaining of this paper is organized as follows. Section II presents the current subjective and objective 3D video quality assessment methods. Section III describes the proposed model behind NVQM in details, and Section IV – the experimental setup. Section V analyses the experimental results and Section V concludes the paper.

II. Related Works

A. Current Objective 3D Video Quality Metrics

Several papers have already investigated objective 3D video quality assessment. [3] and [4] assessed the possiblity of measuring 3D video quality using 2D objective video quality metrics, including PSNR, SSIM and VQM. [3] shows that, by measuring left and right views separately, VQM can effectively predict the overall image quality, and PSNR and SSIM results correlate better with depth perception of 3D video in comparison with VQM. More specific depth map-based stereoscopic video quality assessment is analysed in [4]. In [6], the authors assign a weight of 1/3 of the PSNR score to the left view and the remaining 2/3 to the right view PSNR score. A new perceptual quality metric (PQM) was proposed in [8], which shows better results for 3D video quality in comparison with VQM. This is because it is more sensitive to image degradation and error quantification that happen at pixel level than at sequence level.

In [5], the authors proposed a quality metric that assesses the impact of eye dominance based on spatial frequency by chopping the images into 4 x 4 blocks. Color and sharpness of edge distortion measure (CSED) is proposed in [7], in which the sharpness of edge distortion is considered in depth and color 3D videos.

These 3D video quality metrics have different accuracy levels and advantages. However they all require full reference of the original video source and differ from our proposed no reference network-based metric, which does not require the presence of the original 3D video.

B. Stereoscopic 3D Video Format

The stereoscopic 3D video is composed of a left view and right view video, which can either be stored in one video file or two separate video files. The two offset videos represent the two perspectives of the same scene with a minor deviation (i.e. a human perceives the content with two eyes, and the two videos correspond to the left and right eye of the viewer, respectively). The two views from the videos give the perception of 3D depth while they are combined in the human brain. In a storage format, the two views in stereoscopic 3D video can be top-and-bottom, side-by-side (SBS). While transmitting over the network, the two views are combined into a frame sequential 3D stream, in which the frames are stacked one following another from left view and right view in a frame sequential manner. The details of the techniques can be found in [14].

III. Proposed 3D Video Quality Model

In this section, we firstly present the current 2D video quality assessment model ITU-T G.1070 in details, showing its principles, input parameters and output in form of the mean opinion score (MOS). Secondly, we propose our model based on the G.1070 by considering the effect of depth information of the 3D video.

A. 2D Video Quality Metric using ITU-T G.1070

The ITU-T has standardized a user opinion model for 2D video-telephony applications in G.1070 [15]. This model estimates the 2D video quality in telephony applications by considering the network impairment parameters (i.e. packet loss in video) and encoding parameters, including codec type, video format, key frame interval, and video display size.

The 2D video quality is evaluated by equation (1):

$$V_q = 1 + I_{coding} \, e^{-\frac{Ppl_V}{D_{PplV}}} \qquad (1)$$

where Ppl_V represents packet loss rate, D_{PplV} expresses the degree of video quality robustness due to packet loss, and I_{coding} calculates the basic video quality affected the coding impairment that is introduced by video bit rate (Br_V in kbps) and video frame rate (Fr_V in fps). I_{coding} is calculated as in equation (2):

$$I_{coding} = I_{Ofr} * e^{-\frac{(\ln(Fr_V)-\ln(O_{fr}))^2}{2*D_{FrV}^2}} \qquad (2)$$

In equation (2), parameter O_{fr} represents the optimal video frame rate corresponding to the video bit rate (Br_V) for the best video quality. It is expressed in equation (3):

$$O_{fr} = v_1 + v_2 * Br_v, \ 1 \le O_{fr} \le 30 \quad (3)$$

If $Fr_V = O_{fr}$, then $I_{coding} = I_{Ofr}$. I_{Ofr} is the maximum video quality at the video bit rate and it is calculated by equation (4):

$$I_{Ofr} = v_3 - \frac{v_3}{1 + (\frac{Br_V}{v_4})^{v_5}}, 0 \le I_{Ofr} \le 4 \quad (4)$$

In equation (2), D_{FrV} represents the degree of video quality robustness introduced by frame rate (Fr_V) and is calculated using equation (5) :

$$D_{FrV} = v_6 + v_7 * Br_V, 0 < D_{FrV} \quad (5)$$

At last in equation (1), D_{PplV} represents the degree of video quality robustness due to packet loss rate and is calculated by equation (6) :

$$D_{PplV} = v_{10} + v_{11} * e^{-\frac{Fr_V}{v_8}} + v_{12} * e^{-\frac{Br_V}{v_9}}, 0 < D_{PplV} \quad (6)$$

In the above equations, $v_1, v_2, \ldots, v_{12}$ are derived from 2D subjective video tests and are dependent on the video content bit rate, frame rate, and display size. The methodology for deriving the coefficients in the model is given in the ITU-T G.1070 [8]. As explained in the standard, with the derived coefficients, the related accuracy of the predicted video quality can be evaluated by the Pearson product-moment correlation.

ITU-T G.1070 recommendation includes five sets of coefficients for different display sizes for MPEG-4 and ITU-T H.264, respectively. For the purpose of demonstrating the possibility of extending this model to be applied to 3D video quality assessment, one of the five set of coefficients for MPEG-4 has been used for deriving our model. The derivation of the proposed 3D video quality assessment model is shown in the next sub section.

B. No Reference 3D Video Quality Metric (NVQM)

The proposed 3D video quality model provides an estimated quality by taking into account both packet loss rate and video bit rate. It is non-intrusive compared to other 3D intrusive quality models such as SSIM, VQM, etc.

Unlike the 2D video content, the stereoscopic 3D video consists of views for left and right eyes and the two views work together to provide viewers the three dimensional user perception. The combination of left and right views of stereoscopic 3D video is handled at the display side and uses the information from both left and right views. In a lossy network, the information lost in the left or right view for the same 3D frame (i.e. two views with the same timestamp) might be compensated by the other view, which can increase the overall 3D video quality. If the quality loss is not compensated from the other view, or even worse the lost information affects both views, the displayed 3D frame will be affected, which eventually decreases the overall 3D video quality. Thus we assume that the 3D video quality metric is somehow different from 2D video quality metric when using similar network characteristics.

The 2D video quality model in ITU-T G.1070 provides a good methodology that the bit rate, frame rate and packet loss rate are all considered in equation (1) with further calculations from (2) to (6), and for one display size and codec, the coefficients stay the same.

With the experience given by G.1070, instead of using 1 as the starting point of the MOS result as in (1), coefficient a_1, is used to indicate the different MOS of the perceived 3D video quality. We also denote I_{coding} as a fixed coefficient, expressed by a_2. The factor D_{PplV} representing the robustness to packet loss is kept, but the first two parts of the additive function with coefficients v_{10}, v_{11} and v_8 are combined when substituting Fr_v. This part is represented as a_3. The reason for doing this is that the current model is for a fixed frame rate (i.e. 18 fps) only, being easier to understand and implement.

The proposed no reference 3D Video Quality Metric (NVQM) model is described in equation (7):

$$V_{3Dq} = a_1 + a_2 * e^{-\frac{Ppl_V}{a_3 + a_4 * e^{(-\frac{Br_V}{a_5})}}} \quad (7)$$

In this equation, Ppl_V is the packet loss rate of the 3D video stream in the transmission, and Br_V refers to the 3D video bit rate, including left view and right view. The coefficients are a_1, a_2, a_3, a_4, and a_5, which will be derived from subjective test results.

IV. Experimental Setup

The purpose of the experimental testing is to find the mapping curve between the network characteristics (packet loss rate) and the 3D video perceptual quality, independent from the video content. Firstly different network scenarios with varying network conditions are needed. Secondly, video content independence is ensured by doing the experiment with sufficient number of different 3D video source samples, each with different motion levels.

The video samples are selected from [16]. Five 3D video samples with different content are selected: *dancing*, *kissing*, *running*, *swimming* and *driving* scenarios, which cover high (running and driving), medium (swimming, dance) and low (kiss) motion of the camera relative to the object of interest in the scene. The video clips are around 6 to 14 seconds long according to [17] for the purpose of improving the subjective testing. The network packet loss rate range includes 0%, 0.1%, 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 8% and 10%. All video clips have a fixed frame rate of 18 fps. Each video clip is encoded into two quality levels: high bit-rate as 4 Mbps and low bit-rate as 2 Mbps, both for the combination of left and right view, respectively. In total, there are 2 x 11 x 5 = 110 video samples used in the subjective test. In order to obtain enough samples with a balanced time and human-related costs, each video sample is presented and assessed by 4 different observers.

40 volunteers have participated in the subjective test. Each observer is presented with (110 x 4)/40 = 11 video

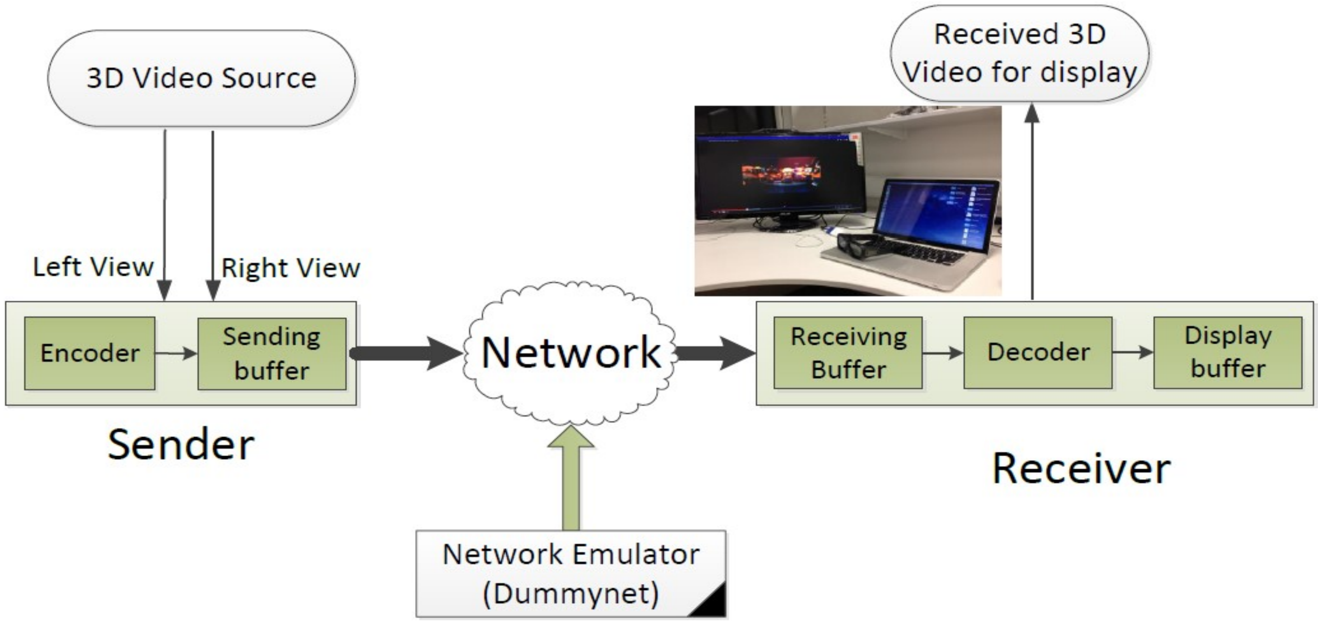

Figure 2. Testing topology

Figure 3. Subjective participant in the testing environment

samples. Within these 11 video samples, the different video content and packet loss in each video are carefully selected so that each observer watches as many different video contents as possible, and the loss rates in these samples cover the cases from low to high with a random order, as well as high and low bit rates. Each participant has received a 5 euro voucher as a thank you gesture.

The testing topology is shown in Figure 2. Different video samples are originally encoded with MPEG-4. The VLC media player[1] is used as sender and receiver on two PCs. At the sender side, VLC is used to encode the left and right video files (in avi format) into RTP streams and send them to the other VLC receiver over the network. A network emulator tool *Dummynet* [18] is used to cause the desired packet loss. The VLC receiver receives the stream sent over the impaired network and decodes the stream to left and right video files in avi format. To acquire the accurate network characteristics, *Wireshark* [19] is used to monitor the received video packets and calculate the real packet loss rate at the receiver side.

In the subjective tests, the video samples are displayed on an ASUS VG278 monitor (27'' display with resolution 1920x1080) with 3D vision 2 support from Nvidia, and the participants wear a pair of the 3D vision 2 wireless active shutter glasses. The viewing distance is set to 1 m as suggested by the monitor manufacturer. The participants are asked to rate these displayed samples from 1 to 5 for overall 3D video quality and 3D depth experience, where 1 indicates the worst and 5 indicates best experience. According to [17], the scores from 1 to 5 mapped to linguistic terms "bad", "poor", "fair", "good", and "excellent", are referred to as the Mean Opinion Score (MOS). The tests were conducted in a 5m x 5m quiet room, while having the display monitor away from direct light from the windows for better viewing experience. A subjective participant in the testing environment is shown in Figure 3.

---

[1] VLC, http://www.videolan.org/index.html

Table 1 Coefficients of NVQM for high & low bit rates

	a_1	a_2	a_3	a_4	a_5
High bit rate (4Mbps)	1.21572	2.49125	-9.85854	44.7371	3000.88
High bit rate Asymptotic Standard Error	+/- 0.5545 (45.61%)	+/- 0.544 (21.84%)	+/- 2.331e+004 (2.364e+005%)	+/- 5.85e+005 (1.308e+006%)	+/- 1.25e+007 (8.332e+005%)
Low bit rate (2Mbps)	1.10136	2.08084	-1.63324	8.33262	3000
Low bit rate Asymptotic Standard Error	+/- 2.601 (236.1%)	+/- 2.544 (122.3%)	+/- 4.483e+006 (2.745e+008%)	+/- 2.751e+007 (3.302e+008%)	+/- 9.793e+009 (6.529e+008%)

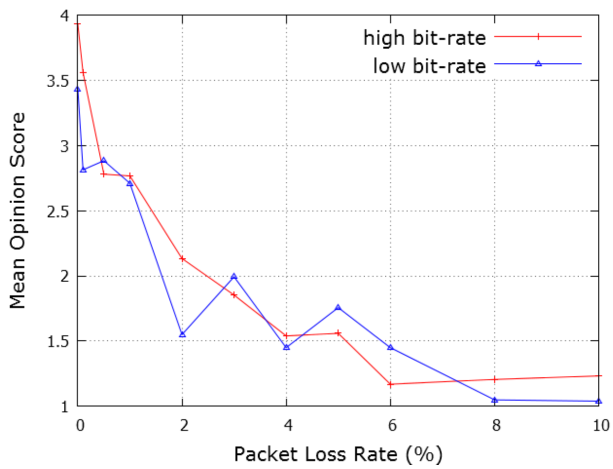

Figure 4. Subjective 3D Mean Opinion Score vs Network loss

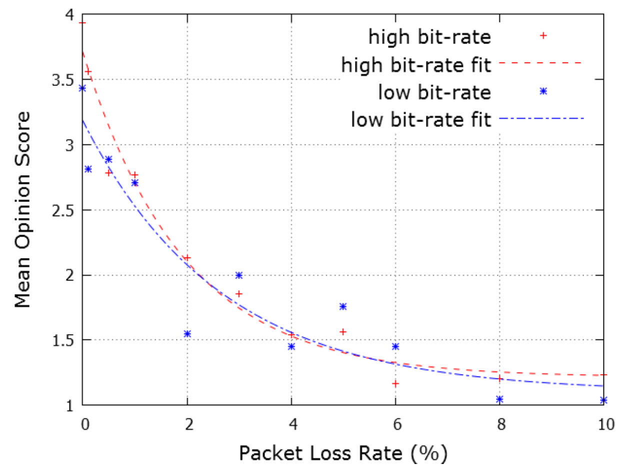

Figure 5. Fitting curve of subjective 3D quality vs Network loss

V. RESULT ANALYSIS

The overall 3D video quality results for the 110 video samples are collected in the subjective testing. Figure 4 shows the subjective MOS values for different samples with different packet loss rates. Each point is an averaged MOS score from 5 samples and each sample is watched repeatedly four times by four different participants. The average MOS score is calculated for each packet loss point from 0% up to 10%. Thus each of the points in the Fig. 4 is an average score of 4 x 5 MOS values. The perceived 3D viewing quality for high bit rate achieves better than low bit rate videos for packet loss rate less than 4%, with a corresponding MOS above 1.5. The MOS level for both high and low bit rates drops rapidly by more than 1.0 when the packet loss rate increases from 0% to 1%, and keep decreasing with similar gradient till 2% packet loss rate. The MOS level for both high and low bit rates videos falls below MOS of 1.5 rapidly after 4% packet loss and the difference between them becomes negligible, which corresponds to unacceptable user experience. Considering all the values, due to the variation of human opinions, the qualities for high bit rate for an increasing packet loss rate maintains a smooth curve, while the qualities for low bit rate varies from around 2% to 5% packet loss rate in a MOS range of 1.5 to 2.0. However this range is known to be associated with bad user quality of experience.

In order to derive a mapping from packet loss and bit rate to the 3D video quality, a fitting curve is calculated and shown in Fig. 5. The two quality fitting curves are derived according to our proposed model in equation (7).

The derived coefficients a_1, a_2, a_3, a_4, and a_5 for equation (7), as well as their standard error, are listed in Table I.

ITU-T G.1070 specifies how to derive the coefficients from *v1*, to *v12*. The basic method uses the combination of multiple bit rates and frame rates under packet loss scenarios, and derives each coefficient while keep one of the variable fixed. The coefficients are approximated based on Least Square Approximation (LSA). In our case, with two varying bit rates, even though 11 packet loss scenarios were create, using LSA will lead to inaccurate results as it is difficult to verify the compromise result for other bit rates. For this reason, we did not use LSA to derive one generic set of confidents, instead we keep two sets of them providing most accurate fitting. Further investigation of more bit-rate levels will be done to derive such a generic function.

In 3D video quality assessment, objective 3D video quality models such as SSIM and VQM use the same weight for both left and right views and simply average the two quality results. The SSIM and VQM scores are computed for each pair of the received/degraded and original reference videos using MSU VQMT [20]. The SSIM results range from 0-1, where larger value indicates better quality and VQM results range from 0 up to around 10 where smaller value means better quality. The two value sets from SSIM and VQM are normalized to 1-5 in MOS scale given in [21] and [22] respectively. The average scores for each packet loss from 5 samples are calculated. By averaging the quality score for two views using G.1070 expressed in equation (1) to (6), the corresponding 3D quality is also obtained. The Pearson Correlation with subjective test results for SSIM, VQM, and G.1070 are listed in Table 2.

Table 2 Pearson Correlation of SSIM, VQM, G.1070, NVQM and Subjective Results

	SSIM	*VQM*	*G.1070*	*NVQM*
High bit rate	0.8712	0.7885	0.9059	0.9736
Low bit rate	0.7474	0.8892	0.8606	0.8961

According to Table 2, it is observed that NVQM outperforms the other 3 objective assessment methods. NVQM has 11% and 19% improvement on accuracy for high and low bit rates compared with SSIM, and 23% higher than VQM for high bit rate. The accuracy improvement for G.1070 for bot bit rates is around 5%. Also, SSIM and G.1070 performs better than VQM in high bit rate while VQM excels SSIM and G.1070 in low bit rate. The difference is caused by the quality variance of the low bit rate 3D videos expected in Figure 4, for the packet loss rate between 2% and 5%. This also explains a relative lower correlation value for low bit rate in the results of SSIM, G.1070 and NVQM. With the Pearson Correlation values, it is confident to say that NVQM correlates well with subjective test results and our metric is reliable under the test environment with 4 Mbps and 2Mbps stereoscopic 3D video in 0% to 10% packet loss.

VI. CONCLUSION & FUTURE WORK

This paper proposes a no reference objective video quality metric (NVQM) for assessing the stereoscopic 3D video quality using network packet loss rate and bit rate as input. NVQM can be used in real time while monitoring network statistics and thus is suitable for decision making process in adaptive 3D video transmissions.

More extensive tests analysis for more bit rates and frame rates are under investigation. Our future work is focused on proposing a more generic metric as an extension of current NVQM.

ACKNOWLEDGMENT

This research was supported in part by the Telecommunications Graduate Initiative (TGI), funded by the Higher Education Authority under the Programme for Research in Third-Level Institutions (PRTLI) Cycle 5 and co-funded under the European Regional Development Fund (ERDF) and in part by Enterprise Ireland Innovation Partnership with Ericsson grant IP/2011/0135.

REFERENCES

[1] Benzie, P.; Watson, J.; Surman, P.; Rakkolainen, I.; Hopf, K.; Urey, H.; Sainov, V.; Von Kopylow, C., "A Survey of 3DTV Displays: Techniques and Technologies," *Circuits and Systems for Video Technology, IEEE Transactions* on , vol.17, no.11, pp.1647-1658, Nov. 2007.

[2] Vetro, A.; Wiegand, T.; Sullivan, G.J., "Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard," *Proceedings of the IEEE* , vol.99, no.4, pp.626-642, Apr. 2011.

[3] Yasakethu, S. L. P., Hewage, C. T., Fernando, W. A. C., and Kondoz, A. M., "Quality analysis for 3D video using 2D video quality models," *IEEE Trans. Consum. Electron.*, vol. 54, no. 4, pp. 1969-1976, Nov. 2008.

[4] Hewage, C. T., Worrall, S. T., Dogan, S., Villette, S., and Kondoz, A. M., "Quality evaluation of color plus depth mapbased stereoscopic video," *IEEE. J. Sel. Topics Signal Process.*, vol. 3, no. 2, pp 304-318, Apr. 2009.

[5] Lu, F., Wang, H., Ji, X., and Er, G., "Quality assessment of 3D asymmetric view coding using spatial frequency dominance model," in *Proc. 3DTV Conference*, Potsdam, Germany. pp. 1-4, May 2009.

[6] Ozbek, N., and Tekalp, A. M., "Unequal inter-view rate allocation using scalable stereo video coding and an objective stereo video quality measure," in *Proc. IEEE Int. Conf. Multimedia and Expo*, pp. 1113-1116, Monterrey, México, Apr. 2008.

[7] Shao, H., Cao, X., and Er, G., "Objective quality of depth image based rendering in 3DTV system," in *Proc. 3DTV Conference*, pp. 1-4, May 2009.

[8] Joveluro, P.; Malekmohamadi, H.; Fernando, W. A C; Kondoz, A.M., "Perceptual Video Quality Metric for 3D video quality assessment," *3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, pp.1-4, Jun. 2010.

[9] Huynh-Thu, Q.; Ghanbari, M., "Scope of validity of PSNR in image/video quality assessment," *Electronics Letters* , vol.44, no.13, pp.800-801, Jun. 2008.

[10] Wang, Z., Lu, L., and Bovik, A. C., "Video quality assessment using structural distortion measurement," *Signal Processing: Image Communication*, vol. 19, no. 2, pp.121 -132, 2004.

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

[12] Lee S.-B.; Muntean, G.-M.; Smeaton, A.F., "Performance-aware replication of distributed pre-recorded IPTV content," *IEEE Transactions on Broadcasting*, vol.55, no.2, pp.516-526, Jun. 2009.

[13] Moller, S.; Wai-Yip Chan; Cote, N.; Falk, T.H.; Raake, A.; Waltermann, Marcel, "Speech Quality Estimation: Models and Trends," *Signal Processing Magazine, IEEE* , vol.28, no.6, pp.18-28, Nov. 2011.

[14] Lang, M., Hornung, A., Wang, O., Poulakos, S., Smolic, A., and Gross, M. (2010). "Nonlinear disparity mapping for stereoscopic 3D," *ACM Transactions on Graphics (TOG)*, vol. 29, issue 4, no.75, Jul. 2010.

[15] ITU-T G.1070", "Opinion model for video-telephony applications", Jul. 2007.

[16] Hadfield, S.; Bowden, R., "Hollywood 3D: Recognizing Actions in 3D Natural Scenes," *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference*, vol., no., pp.3398,3405, 23-28, Jun. 2013.

[17] "ITU-R BT.500-13", Methodology for the subjective assessment of the quality of television pictures, Jan. 2012.

[18] Carbone, M., and Rizzo, L., "Dummynet revisited," *ACM SIGCOMM Computer Communication Review*, vol. 40, no.2, pp.12-20, Apr. 2010.

[19] Orebaugh, A., Ramirez, G., and Beale, J. (2006). "Wireshark & Ethereal network protocol analyzer toolkit," Syngress, 2006.

[20] MSU Video Quality Measurement Tool. Available in: http://compression.ru/video/quality_measure/video_measurement_tool_en.htm.

[21] Zinner, T., Abboud, O., Hohlfeld, O., Hossfeld, T., and Tran-Gia, P. (2010, March). "Towards qoe management for scalable video streaming," *21th ITC Specialist Seminar on Multimedia Applications-Traffic, Performance and QoE* , pp. 64-69, Mar. 2010.

[22] Dymarski, P., Kula, S., and Huy, T. N. (2011). "QoS Conditions for VoIP and VoD," *Journal of Telecommunications & Information Technology*, vol. 2011, no.3, pp.29-37, Mar. 2011.