2016

Hybrid Sequencing of Uncompressed and Compressed 3D Stereoscopic Video: A Preliminary Quality Evaluation

Siu Ming, Vincent Choy
Hong Kong Institute of Vocational Education (Tuen Mun), Vocational Training Council, vchoy@vtc.edu.hk

Eva Cheng
RMIT University, ecc04@uow.edu.au

Ian Burnett
University of Technology, Sydney, ian.burnett@uts.edu.au

Hybrid Sequencing of Uncompressed and Compressed 3D Stereoscopic Video: A Preliminary Quality Evaluation

Siu-Ming CHOY[1, 2], Eva CHENG[1], Ian BURNETT[3]

[1] School of Electrical and Computer Engineering,
RMIT University, Melbourne, Australia
[2] Department of Engineering, Hong Kong Institute of Vocational Education,
Vocational Training Council, Hong Kong
[3] Faculty of Engineering and Information Technology
University of Technology Sydney, Australia

Abstract

The use of 3D stereoscopic technology with high quality videos can provide visual entertainment to viewers. However, the bandwidth of typical communication channels cannot transmit uncompressed 3D videos, resulting in the need for video quality compression. This paper presents a series of preliminary studies to investigate the subjective perception of uncompressed and compressed video sequences, and proposes the 'hybrid' sequencing of uncompressed and compressed content in a single stereoscopic 3D video as an alternative approach for limited bandwidth transmission. However, the hybrid uncompressed/compressed sequencing of stereoscopic 3D video may affect the correlation between the left and right views of the stereoscopic videos required for depth perception, potentially leading to lower Quality of Experience (QoE) of viewers. This paper therefore investigates both the objective and subjective quality evaluation of the proposed hybrid sequencing of stereoscopic video sequences. Initial investigations into objective metrics to measure the difference in quality of the two stereoscopic views due to the proposed hybrid sequencing of uncompressed and compressed videos were also conducted.

Keywords: 3D, stereoscopic video, PSNR, QoE, subjective video assessment

1 INTRODUCTION

Stereoscopic 3D video has been widely popular in many commercial markets and the entertainment industry. Higher resolution stereoscopic videos, such as Full-HD and 4K, are available on the commercial market in high-budget movies and also 3D-enabled consumer TVs. Despite the increasing need for higher compatibility of stereoscopic videos with other devices (such as mobile devices for gaming etc.), it is not always feasible to transmit uncompressed or very high quality 3D videos through communication channels due to the limited bitrate transmission.

Researchers have proposed several techniques for the compression of stereoscopic 3D video content for transmission: Vetro et al. [1] compared different 3D representation formats and coding architectures to evaluate the performance of various 3D video compression approaches, and Yao et al. [2] suggested 3D video coding algorithms to distribute channel bandwidth dynamically. However, the bitrate for video coding approaches using scalable stereo video coding to control the bitrate to limit bandwidth resulted in higher computational complexity [3].

Pinson et al. [4] recently proposed that the video coding difficulty of 'hard-to-code' 2D sequences resulted in lower quality than the 'easy-to-code' sequences. An objective complexity metric was thus recommended to evaluate the scene complexity, but the metric of subjective tests was not addressed. Tominaga et al. [5] compared different subjective assessment methods, such as Double-Stimulus Continuous Quality Scale (DSCQS), Absolute Category Rating (ACR) and Degradation Category Rating (DCR). Further, Seo et al. [6] found traditional video evaluation methods to be inappropriate for assessment of stereoscopic video, and proposed a new video quality metric for compressed stereoscopic video. Kawano et al. [7] compared different subjective assessment methods for 2D and 3D video quality and suggested that the ACR method was the most suitable to assess participants' stability and assessment time of 3D videos. Further studies using the ACR method could identify the characteristics of different 3D videos of varying video quality.

As an alternative to the computationally complex scalable video coding techniques and to cater for stereoscopic videos with scenes that may be susceptible to potential artifacts introduced into the left/right views from compression, this paper proposes the 'hybrid' sequencing of uncompressed and compressed content in a single stereoscopic 3D video as an alternative approach for limited bandwidth transmission. A series of preliminary studies are conducted using the ACR method to investigate the subjective perception of uncompressed and compressed video sequences using the proposed hybrid sequencing. In addition, whether the proposed hybrid uncompressed/compressed sequencing of stereoscopic 3D video affects the correlation between the left and right views of the stereoscopic videos required for depth perception is investigated using objective metrics.

The paper is organized as follows: Section 2 discusses the background for the objective and subjective evaluations conducted in this paper, and Section 3 details the video quality experiments. Section 4 reviews and discusses the experimental results obtained and Section 5 concludes the paper.

2 BACKGROUND

To study the objective and subjective evaluation of 3D videos, a number of measurement assessment methods were evaluated in this paper as follows:

2.1 PSNR

PSNR (Peak Signal-to-Noise Ratio) is a commonly utilized method to measure objective evaluation and assess the correlation between an uncompressed image and the compressed image. The PSNR is calculated as follows:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left|S(i,j)-C(i,j)\right|^2$$
$$PSNR = 10\log_{10}\left(\frac{MAX_I^2}{MSE}\right) \quad (1)$$

Where $S(i, j)$ is the uncompressed image, $C(i, j)$ is the compressed image and MAX_I is the maximum value of a pixel. If the pixels are represented in 8 bits per pixel, the MAX_I value is 255.

2.2 Subjective quality assessment

ITU (The International Telecommunication Union) standardized a series of subjective methods to assess video quality: ITU-T P.910 [8] documents the subjective video quality assessment methods for multimedia applications. The ACR method, shown in Figure 1, was adopted for the subjective quality assessments conducted in this paper. That is,

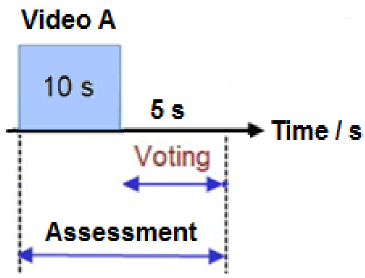


Figure 1. ACR method

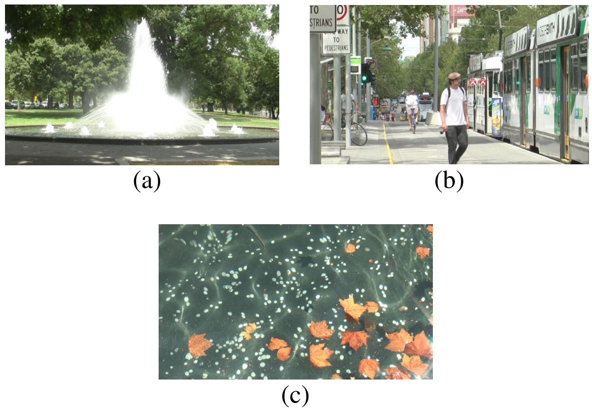


Figure 2. Videos selected from the RMIT3DV database: (a) 'Water fountain', (b) 'Tram stop', (c) 'Wishing well'.

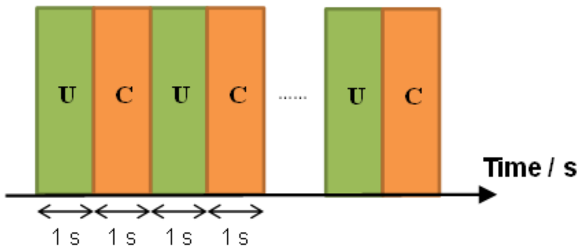


Figure 3. The orientation of the proposed hybrid video sequencing, where '**U**' and '**C**' represent uncompressed and compressed video frames, respectively

the video under test was presented once and subsequently ranked by the participant afterwards.

3 PROPOSED METHODOLOGY

Three stereoscopic 3D videos were selected from the RMIT3DV [9] database, as shown in Figure 2. The characteristics of the videos are detailed in Table 1. All videos were recorded in 1080p HD resolution, uncompressed 10-bit, and YUV 4:2:2 at 25 fps, stored in MOV format.

Table 1. Characteristics of database video sequences

Name of database video sequences	Characteristics
Water fountain	Difficult for encoding
Tram stop	Easy for encoding
Wishing well	Many small objects

Table 2. PSNR comparison for database video sequences

PSNR(dB) / *Video quality*	Water fountain	Tram stop	Wishing well
1080p	35.78	35.32	35.12
720p	36.00	35.87	35.23
*1080p uncompressed	17.52	12.35	16.34
*1080p hybrid	17.54	12.36	16.37
*1080p compressed	17.57	12.37	16.40
*720p uncompressed	17.81	12.44	16.47
*720p compressed	17.84	12.45	16.52

Table 3. Video sequence pairs for subjective assessment

Video Sequence pairs	Video 1	Video 2
1	Compressed 1080p	Compressed 720p
2	Compressed 720p	Compressed 1080p
3	Uncompressed 1080p	Hybrid 1080p
4	Uncompressed 720p	Hybrid 1080p

Different versions of the stereoscopic video sequences were then generated with FFmpeg software [10] for the subjective evaluations conducted in this paper: 1) original uncompressed video in MOV format; 2) the proposed hybrid sequencing, which is generated by the alternative sequencing of one second each of uncompressed and compressed sequences within each video as shown in Figure 3 (also stored in MOV format); 3) compressed video coded by AVC/H.264 codec in MP4 format. The video formats of uncompressed and compressed videos were also generated in 720p and 1080p resolutions to investigate the effect of varying frame resolution.

3.1. Objective quality assessment

For objective quality assessment, the PSNR values were computed to compare between uncompressed, compressed, and the proposed hybrid sequencing at the same video frame resolution. The PSNR was also calculated between the left and right videos in order to detect the differences (if any) in the stereoscopic disparity information. The calculation of the PSNR is shown in Table 2, where '*' denotes the computation between the left and right videos.

3.2. Subjective quality assessment

For subjective evaluations, a series of 3D videos were displayed to the participants. For the experiments conducted in this paper, ten participants (8 male and 2 female) aged from 20 to 38 years old. Prior to the experiment, participants completed 3D vision tests based on ITU-R BT.2021 standard [11].

A) Experimental apparatus and conditions

In the subjective quality assessments, all stereoscopic 3D videos were presented on a 25.5'' Panasonic BT-3DL2550 Full HD LCD 3D monitor. Participants were required to wear passive 3D glasses throughout the experiment.

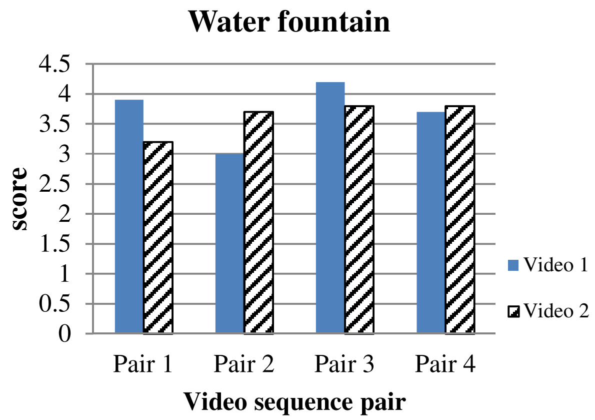
B) Experimental procedures

The THX Cinema Certification specification was utilised in the subjective quality assessments, in which a 36 degree viewing angle and 0.9m viewing distance was adopted in the experiments of this paper [12]. As pairwise comparison tests were conducted, each participant was asked to sit in front of the 3D monitor to watch the first 3D video for 10 seconds. The participants were then required to rank the subjective video quality using the five-point ACR scale (1: bad, 2: poor, 3: fair, 4: good, 5: excellent) in 10 seconds. Afterwards, the participant was required to watch the second 3D video (of a different quality) for 10 seconds and asked to rank the video quality again. Each participant was required to watch the three videos as selected from the RMIT3DV database (Figure 2), where a combination of four different video pairs was presented for comparison for each video (as summarised in Table 3). The order of the four pair-wise video tests were randomly chosen to minimize personal bias. Thus, in total, 12 video pairs were used for the experiments.

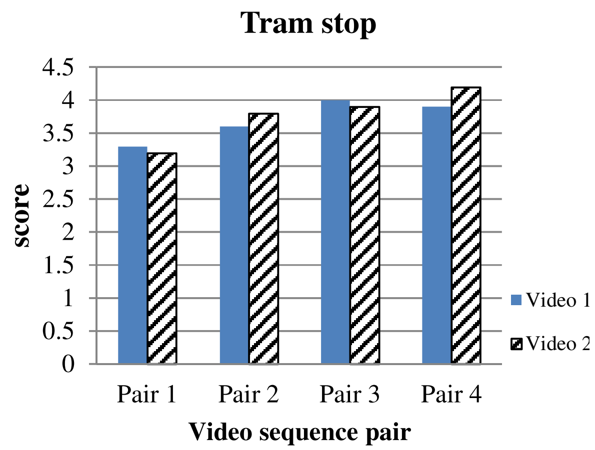
4 DISCUSSIONS OF EXPERIMETAL RESULTS

4.1 PSNR of 3D videos

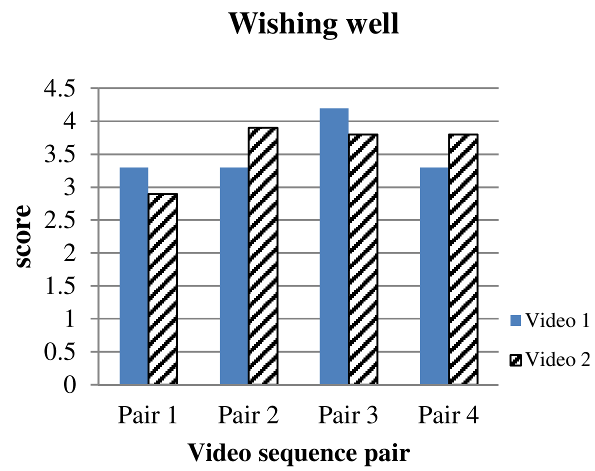
For the PSNR comparison in Table 2, the overall PSNR values of 1080p resolution among all video sequences were lower than for 720p resolution. The results revealed that higher resolutions may result in more error when generated from uncompressed to compressed video, potentially leading to lower video quality at higher resolutions although using the same video codec. In addition, the PSNR values between the left and right videos were compared to investigate the relationship between objective and subjective video quality evaluation [6]. Among the left and right videos tested, the 'Tram stop' video obtained the lowest PSNR values while 'Water fountain' obtained the highest PSNR values. However,



(a) The video sequence: Water fountain



(b) The video sequence: Tram stop



(c) The video sequence: Wishing well

Figure 4. Average score distributions of the ACR method for three stereoscopic 3D videos tested

the subjective evaluations suggested that the 'Tram stop' video demonstrated the least score distribution variation amongst participants in response to the various quality and resolutions tested. These preliminary results suggested that the disparity of stereoscopic 3D videos may not be the only factor that affects 3D video quality. The subjective evaluations indicated that other characteristics of 3D videos, such as the motion in the scene, encoding method and the picture orientation could also affect the perceptual video quality.

4.2 ACR subjective quality assessments

The average ACR scores for each subjective pairwise comparison of varying video quality is shown in Figure 4. In the video pairs 1 and 2 shown in Figure 4, the aim of this series of subjective evaluations was to investigate whether the ACR scores varied between videos of low (high) to high (low) resolution transition. From the preliminary subjective evaluations conducted, pairs 1 and 2 in Figure 4 suggest that the ACR scores from the 'Tram stop' video exhibited the least variation between different video qualities tested, possibly due to the slow movement of both pedestrians and the tram in the video.

The pairwise comparison of pairs 3 and 4 aimed to compare the uncompressed video to the proposed hybrid video sequencing (as shown in Figure 3). From Figure 4, it can be seen for the pair 3 test that the overall ranking of the hybrid video was lower than the uncompressed video when at the same resolution (1080p). However, when compared to a lower resolution of the uncompressed video (720p) in the pair 4 test, the proposed hybrid sequencing results in higher ACR ratings for all the three videos tested. Further, similar to the pair 1 and 2 tests, for the pair 3 and 4 tests the ACR score of the 'Tram stop' video again exhibited the least score variation across the pairwise tests. The less stability in ACR scores for the 'Water fountain' video could be due to the video containing fast movement of water inducing a high 3D effect with over-distorted images. In contrast, the less stable ACR scores from the 'Wishing well' video could be due to the video containing many small objects with time-varying depth perception, which could attract more participants to concentrate on particular objects to potentially lead to participants' having higher sensitivity to 3D quality when comparing between different video qualities.

5 CONCLUSION

The preliminary objective and subjective experiments reported in this paper investigated the feasibility of the proposed hybrid sequencing of uncompressed and compressed stereoscopic 3D video, studying how the sequencing may affect the perceptual video quality as subjectively rated by viewers. The results suggest that the proposed hybrid sequencing may result in a similar perceptual quality to uncompressed video with 'stable' 3D scenes and scenes with moderate 3D effects. However, immediate future work will conduct subjective tests with a larger sample of viewers to further verify the preliminary results obtained in this paper, also investigating higher compression qualities to compare with the proposed hybrid sequencing. Further analysis into which 'types' of 3D scenes, the scene complexity and level of 3D effect which is suited to the proposed hybrid sequencing will also be conducted.

REFERENCES

[1] A. Vetro, A. M. Tourapis, K. Muller, and T. Chen, "3D-TV Content Storage and Transmission," *IEEE Transactions on Broadcasting,* vol. 57, pp. 384-394, 2011.

[2] W. Yao, L. P. Chau, and S. Rahardja, "Joint Rate Allocation of Stereoscopic 3D Videos in Next-Generation Broadcast Applications," *IEEE Transactions on Broadcasting,* vol. 59, pp. 445-454, 2013.

[3] Z. Chen, W. Lin, and K. N. Ngan, "Perceptual video coding: Challenges and approaches," in *Multimedia and Expo (ICME), 2010 IEEE International Conference on*, 2010, pp. 784-789.

[4] M. H. Pinson and M. Barkowsky, "Selecting sc enes for 2D and 3D subjective video quality tests," *Journal on Image and Video Processing 2013,* 2013.

[5] T. Tominaga, T. Hayashi, J. Okamoto, and A. Takahashi, "Performance comparisons of subjective quality assessment methods for mobile video," in *Quality of Multimedia Experience (QoMEX), 2010 Second International Workshop on*, 2010, pp. 82-87.

[6] J. Seo, X. Liu, D. Kim, and K. Sohn, "An Objective Video Quality Metric for Compressed Stereoscopic Video," *Circuits, Systems, and Signal Processing,* vol. 31, pp. 1089-1107, 2012/06/01 2012.

[7] T. Kawano, K. Yamagishi, and T. Hayashi, "Performance comparison of subjective assessment methods for 3D video quality," in *Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on*, 2012, pp. 218-223.

[8] I.-T. R. P.910, "Subjective video quality assessment methods for multimedia applications," ed, 2008.

[9] E. Cheng, P. Burton, J. Burton, A. Joseski, and I. Burnett, "RMIT3DV: Pre-announcement of a creative commons uncompressed HD 3D video database," in *Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on*, 2012, pp. 212-217.

[10] FFmpeg. Available: https://www.ffmpeg.org/

[11] A. S. Umar, R. M. Swash, and A. H. Sadka, "Subjective quality assessment of 3D videos," Victoria Falls, Livingstone, 2011.

[12] S.-M. Choy, K.-H. Chiu, E. Cheng, and I. Burnett, "3D fatigue from stereoscopic 3D video displays: Comparing objective and subjective tests using electroencephalography," in *TENCON 2015 - 2015 IEEE Region 10 Conference*, 2015, pp. 1-4.

Contact E-mail Address:
siumingvincent.choy@rmit.edu.au