Subjective Evaluation of the Effects of Packet Loss on HEVC Encoded Video Streams

James Nightingale, Qi Wang, Christos Grecos
Centre for Audio-Visual Communications and Networks
University of the West of Scotland
Paisley, United Kingdom
Email {firstname.lastname@uws.ac.uk}

Sergio Goma
Multimedia R&D
Qualcomm Inc.
San Diego, USA
Email {sgoma@qti.qualcomm.com}

***Abstract*— The emerging High Efficiency Video Coding standard (HEVC) will bring the benefit of delivering the same statistical quality at about half of the bandwidths required in the current H.264/AVC standard. Such significantly higher compression efficiency of HEVC will, however, potentially lead to higher sensitivity to packet loss and thus have a great impact on the users of portable consumer devices such as smartphones and tablets when delivering HEVC encoded video over loss-prone networks, thereby adversely affecting the user's quality of experience (QoE). Existing subjective evaluations of the perceptual quality of HEVC have focused on its performance in loss-free environments. In this work, we empirically transmit HEVC streams over a hybrid wired/wireless network at typical (UK) mobile broadband speeds under a range of packet loss conditions, using typical smartphone and tablet resolutions. Our subjective evaluation experiments quantify the effect, on perceptual quality, of packet loss in HEVC streams and establish a packet loss rate threshold of 3% beyond which users find poor perceptual quality has a detrimental effect on their QoE. Furthermore, we employ two error concealment schemes to mitigate the impact of packet loss/corruption and investigate their effectiveness on users' QoE.**

***Index Terms*— HEVC, subjective evaluation, perceptual quality, QoE, packet loss.**

I. INTRODUCTION

Previous subjective evaluations of HEVC were primarily conducted as part of the standardization effort for HEVC. They aimed to establish if the objective of reducing bitrates by 50% (from H.264/AVC) while maintaining perceptual quality had been met. These studies, one example of which was reported in [2] all compared the initial raw video directly with the output from an HEVC decoder. These studies assumed a lossless delivery of an HEVC encoded stream to the decoder. No consideration was given to any losses which may be incurred during transmission. Another subjective evaluation by Garcia and Kalva [3] questioned the value of HEVC encoding in the context of a smartphone by concluding that there was little perceptual difference between HEVC and H.264/AVC at typical smartphone resolutions and bandwidths. However their study only used a single bandwidth/spatial resolution combination and did not consider either the bandwidth saving that may be achieved (at the same perceptual quality) or the potential for transmission losses in their smartphone scenario.

This work evaluates the perceptual effect of packet loss on HEVC streams and establishes that packet loss in excess of 3% leads to a significant drop in viewer satisfaction. We also show that viewers are less tolerant of visual artifacts produced during error concealment than the 'freezing' caused by the use of a basic frame copy error concealment scheme. The mean opinion scores (MOS) from our subjective study are supplemented by comparison with the peak signal to noise ratio (PSNR), and video quality (VQM) metrics.

II. SUBJECTIVE EVALUATION

In our evaluation we employed 16 human subjects aged between 22 and 50, all were in good health and wore appropriate corrective lenses where required. Sequences were displayed against a mid-grey background on a Hewlett Packard 22" LCD monitor with native resolution of 1920*1080. Viewers sat upright in at the distance recommended for each spatial resolution in [5]. We used the double stimulus impairment scale (DSIS) defined in [5] where the subject was presented with test sequences in pairs. The first sequence was the reference video followed by the impaired video. They were asked to score the impaired sequences on a scale from 1 to 5 as defined in [5] where 5 is imperceptible difference and 1 is very annoying. Six video HEVC test sequences (*Racehorses, Flowervase, Blowingbubbles, Basketballpass, Basketballdrilltext, Keiba*) were each encoded using the HEVC reference software (version HM 8.0) [1] at two different spatial resolutions (416*240 and 832*480) representative of the resolutions of mid range smartphones and tablet devices respectively. Three bandwidth testing points of 500Kbps, 1400Kbps and 2200Kbps were used which represent low, medium and high bandwidth values for typical UK mobile broadband. The quantization parameter (QP) was varied to produce bitstreams closely matched to the three target bitrates. Some streams were encoded using the HM 8.0 [1] default setting where each encoded slice contained exactly one picture and was encapsulated in a single NAL unit. The remaining sequences were encoded with a NAL unit size that matched the maximum transmission unit (MTU) of 1500 bytes used within our testbed resulting in multiple NAL units for each picture. Each test sequence was transmitted over a hybrid wired/wireless testbed using the HEVStream framework [4]

978-1-4799-1412-8/13/$31.00 ©2013 IEEE

with packet loss rates of 0%, 1%, 3% and 5% applied. Streams were decoded using a version of the HM 8.0 [1] reference decoder that had been modified to compensate for NAL units contained in packets lost during transmission. Two types of basic error concealment were adopted, in the first method an entire picture is copied to the current position from the nearest picture (temporally) in the reference picture set. The second method only copies the lost/damaged area (co-located blocks) from the nearest reference picture to the current picture. The first method was applied when an entire picture was lost or the first NAL unit of a picture was lost. The second method was applied when a slice containing only part of a picture was lost.

III. Results

Fig. 1 presents the mean opinion scores for the *Racehorses* video sequence at WQVGA (416*240) resolution. Mean opinion score is presented on a scale from 1 to 10, where 10 is the highest possible value. The MOS value shown is the average MOS with the error bars representing the left confidence border (lower) and right confidence border (upper) as defined in [5]. The figure shows that perceived quality is high (average MOS = 9.17) at all of the bandwidths tested when no packet loss occurs, dropping to 8.86, 6.58 and 5.31 at 1%, 3% and 5% packet loss respectively.

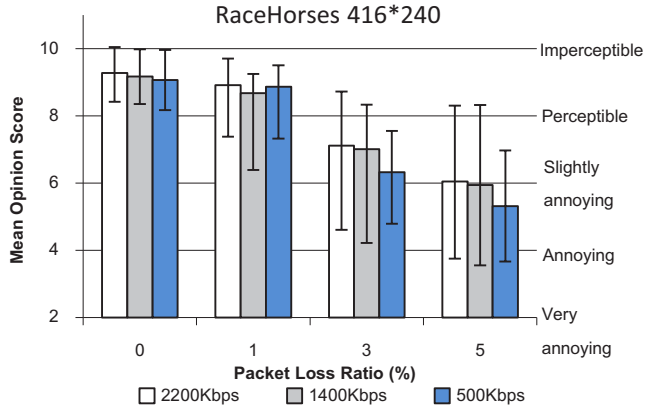


Fig. 1. Mean opinion scores for the WQVGA *Racehorses* sequence.

Average MOS results from all of the video sequence/testing point combinations used are compared to the PSNR values in Fig. 2. From this figure it can be seen that MOS values (viewer satisfaction) decline at a greater rate than PSNR values as packet loss increases. Variability in subjective scores was observed to increase with packet loss and is shown as the standard deviation in MOS in Fig. 2. VQM increased with packet loss (lower is better). Our study has shown that the relationship between packet loss and MOS is a complex one. We have noted that several other factors act in combination with packet loss to contribute to the viewer's perception of quality. For example viewers were more tolerant (by an average MOS of 0.98) of the error concealment method that employed whole picture copy than the one that copied only the missing area of a picture. This observation has implications for both error concealment in HEVC and the sizing choices for slices/NAL units. Although the MOS degradation depicted in Fig. 2 generally held true for all experiments, there were variations between sequences which we attribute to the content of the video clip. There were also variations between spatial resolutions of the same sequence.

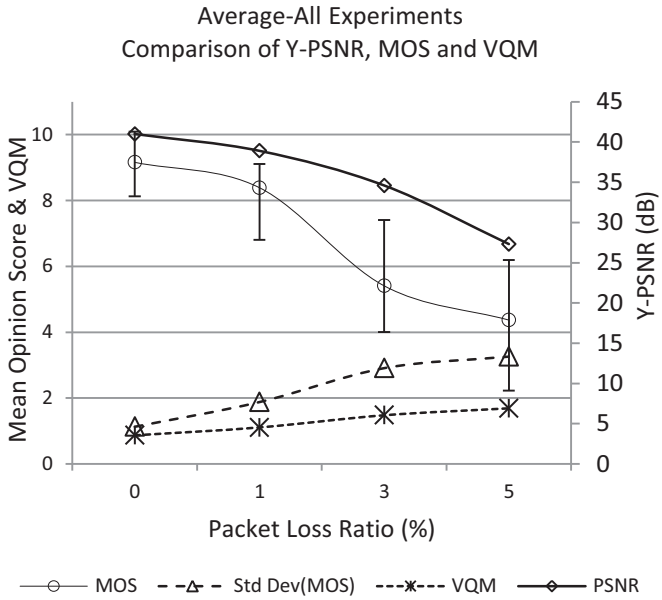


Fig. 2 Comparison of average MOS, VQM and PSNR across all experiments

IV. Conclusions

Our study has determined that human perception of video quality in HEVC streams is significantly affected by packet loss during transmission. At 1% packet loss the average mean opinion score (MOS) drops from 9.16 to 8.39 indicating that the packet loss is perceptible but not detrimental to the viewer experience. Whereas a packet loss ratio of 3% reduces the average MOS to 5.41 and a 5% loss further reduces it to 4.37. We therefore conclude that 3% packet loss represents the threshold of user tolerance where sequences were typically scored in the 'slightly annoying' category. We also observed that the variability of MOS scores rose as packet loss increased indicating a range of opinions on the level of quality degradation encountered. Finally we have shown that the choice of error concealment method employed can significantly influence viewer satisfaction.

References

[1] JCT-VC, "HEVC Test model under consideration (TMuC) Revision 8.0," *JCT-VC Contribution* HM 8.0, Nov. 2012.

[2] J. Ohm et al, "Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC)," *Circuits and Systems for Video Technology, IEEE Transactions on* , vol.22, no.12, pp.1669-1684, Dec. 2012.

[3] R. Garcia and H. Kalva, "Human mobile-device interaction on HEVC and H.264 subjective evaluation for video use in mobile environment," *Proc. of IEEE International Conference on Consumer Electronics (ICCE 2013),* Jan. 2013.

[4] J. Nightingale, Q. Wang, and C. Grecos, "HEVStream: a framework for streaming and evaluation of high efficiency video coding (HEVC) content in loss-prone networks," *Consumer Electronics, IEEE Transactions on*, vol.58, no.2, pp.404-412, May 2012.

[5] ITU, "Methodology for the subjective assessment of television pictures," ITU-R BT500-13, Jan. 2012.