Packet Loss Visibility Across SD, HD, 3D, and UHD Video Streams

A.O. Adeyemi-Ejeye, M. Alreshoodi, L. Al-Jobouri, M. Fleury, J. Woods

Abstract

The trend towards video streaming with increased spatial resolutions and dimensions, SD, HD, 3D, and 4kUHD, even for portable devices has important implications for displayed video quality. There is an interplay between packetization, packet loss visibility, choice of codec, and viewing conditions, which implies that prior studies at lower resolutions may not be as relevant. This paper presents two sets of experiments, the one at a Variable BitRate (VBR) and the other at a Constant BitRate(CBR), which highlight different aspects of the interpretation. The latter experiments also compare and contrast encoding with either an H.264 or an High Efficiency Video Coding (HEVC) codec, with all results recorded as objective Mean Opinion Score (MOS). The video quality assessments will be of interest to those considering: the bitrates and expected quality in error-prone environments; or, in fact, whether to use a reliable transport protocol to prevent all errors, at a cost in jitter and latency, rather than tolerate low levels of packet errors.

Keywords: Video streaming, packet loss visibility, objective MOS, Beyond HD

1. Introduction

Packet Loss Visibility (PLV) [1][2] appraises video quality according to the network response and in doing so may cut across assessments based on the bitrate and/or compression ratio. For example, we have found that when packet loss is taken into account, video compressed at a lower quantization parameter (QP), which would normally result in a higher video quality may result in lower-quality video than video compressed with a higher QP. This assumes that the packetization structures are the same for both QPs and that 'video quality' refers to objectified Mean Opinion Score (MOS) [3]. (Objectified MOS, herein, results from a direct mapping from either the Structural

SIMilarity (SSIM) index [4] or the Video Quality Metric (VQM) [5].) In fact, PLV also may change in response to variations in: bitrate; content type; and codec [6]. However, in this paper, the main focus is upon the relationship between spatial-resolution and PLV, because the current trend is towards video streaming at ever higher resolutions, including HD (1280×720 pixels/frame progressive (p) scanning), 3D (both Standard Definition (SD) (480×720p) and HD), and 4k Ultra High Definition (UHD)(4096×2160p)[7].

That trend extends to portable devices, which must operate in error-prone wireless environments subject to packet loss. System-on-chips have been designed [8] for HD video on smart phones and mobile 3D video has attracted researchers [9]. Indeed, HD video has a $3.1 \times$ picture height for an ideal viewing distance compared to $7.1 \times$ picture height for SD [10], making HD video particularly appropriate for portable displays. (For example, at a distance of 3.1 multiplied by the HD picture height the scan lines become invisible, whereas at a closer distance a viewer can distinguish between them. Viewers tend to select that viewing distance at which the scan lines first become invisible.) The video plus depth format for 3D TV [8] is also appropriate for portable devices, as the main obstacle to auto-stereoscopic (glasses-free) displays [11] is the need for a restricted viewing angle, which is already present on portable devices, as they are rarely used for social viewing by more than one viewer. Furthermore, UHD broadcast transmission to portable devices has already been demonstrated at the 2014 Broadcast Asia conference employing Digital Video Broadcast terrestrial second-generation (DVB-T2) transmission. Compression was through an High Efficiency Video Coding (HEVC) standard codec, which was designed with higher resolutions in mind [12], because, as is well-known, it achieves up to 50% bitrate savings over the prior H.264/Advanced Video Coding (AVC) standard. However, as this paper details, there also may be a cost in terms of exposure to packet loss, and in commercial terms there is also a risk if a service to portable devices is not appreciated by its audience.

Thus, user expectations of mobile displays extend up to broadcast quality owing to both the apparent sophistication of smart phones and the ability to offload high-intensity computation to cloud processing [13]. Despite those expectations, the reality is that video is still streamed to such devices over error-prone wireless networks, where it is impossible to make 'less error-prone' by altering the physical channel. It is true that TCP-based pseudo-streaming with some form of HTTP Adaptive Streaming (HAS) [14] can be applied to wireless communication. When an error occurs packets are retransmitted.

However, though there are no errors present at the receiver, large buffers are required, implying start-up delay affecting short video clips, and, for other than short video sequences, there is a risk of increased and annoying jitter. The impact of jitter is particular noticeable in interactive services such as mobile video gaming, teleconferencing, and telemedicine. Alternatively, if an Internet Protocol (IP)/User Datagram Protocol (UDP)/Real-time Transport Protocol (RTP) variety of streaming is employed [15] then packet loss can impact upon the receiver.

This paper also investigates another form of video transmission over a network, namely by packing MPEG-2 Transport Stream (TS) media container packets (maximum size 188 B) directly into UDP network packets. Though this paper studies PLV rather than jitter, all three forms of streaming (HAS, IP/UDP/RTP and MPEG-2 TS/UDP) run the risk of complaints published on the Web, which can escalate in their impact [16]. Therefore, it becomes important to study PLV because higher spatial resolutions are becoming more common.

However, most video quality evaluation studies were concerned with video resolutions below or occasionally up to HD and not beyond that, see [17] [18]. Nevertheless, there is indirect evidence that changes in coding format result in differing responses to packet loss. Thus, [6] showed that H.264/AVC quality drops dramatically for even low packet loss rates (0.02%), while MPEG-2 quality drops by much less. The implication is that for equivalent transmitter HD quality the older codec achieves a better receiver quality once packet loss is taken into account. It is also important to notice that Pinson et al. [6] did this with experiments that adjusted for the expected coding gain between MPEG-2 and H.264/AVC to allow the PLV of video of expected equal quality to be compared. Thus, for MPEG-2 a bitrate of 6 Mbps was compared with an H.264/AVC bit-rate of 2 Mbps amongst others. The same approach is also taken in [19] in that bitrate savings between codecs are compared across equal objective video quality. We have, thus, selected bitrates that adjust for expected coding gains to achieve a comparison across equal quality video. To the best of our knowledge, the current work is the first to compare HD and 4kUHD video quality in respect to PLV for H.264/AVC and HEVC codecs.

Comparing between resolutions rather than between codecs, contrasting views emerge. [1] noted that packet loss was more visible at HD than at SD, because HDTV occupies a larger field of view. However, [2] suggested that the quality degradation as a result of a packet loss in HD is much lower than in SD because the relative amount of information carried in that packet is smaller. As a result, a lost packet affects less macroblocks (MBs) in HD, and, therefore, causes less spatio-temporal error propagation. Consequently, it is incumbent upon researchers to establish how packet structure affects video quality, irrespective of differences in bitrate, especially at 3D and UHD resolutions.

Ideally, subjective assessment is required to assess video quality. Unfortunately, managers of video streaming normally do not have access to a panel of viewers [20] [21], owing principally to: time restrictions; and the difficulty of assembling a suitable set of viewers. Subjective testing also does not allow a real-time response to changes in packet loss rate (PLR) or packet structures and is not repeatable. However, objective subjective ratings can approximate the results of subjective testing with a high degree of correlation. For example, the Video Quality Experts Group (VQEG) Full-Reference Television (FR-TV) Phase II tests for VQM [22] resulted in Pearson linear correlation coefficients (PCCs) with Difference MOS (DMOS) from subjective tests, of above 0.9 [5]. Following the VQEG evaluation procedure SSIM also resulted in correlations well above 0.9 and in [4] outperformed four other models on the authors' still image database. Since the original SSIM presentation [23], a number of refinements have also occurred. These give good confidence that SSIM along with VQM are excellent objective measures of MOS. The result is that this paper has been able to combine experimental results originally measured by either SSIM or VQM in one common objective MOS score. The results are repeatable and are arrived at in a practical manner.

The contribution of this paper is to examine the relative impact of packet loss upon PLV across SD, HD, 3D, and 4KUHD video streams. Examining PLV rather than compression gain gives a realistic guide to what can be expected from higher resolutions in error prone environments, especially for portable devices. Realistic video configurations have, thus, been used and the important influence of content type, bitrate, and packetization is exposed. Because VQM and SSIM have been employed, the link between: packet-loss rate; error burst length; and content type has been directly mapped to MOS subjective ratings. In this way, the paper exposes the relationship between packetization and video resolution.

As a motivating argument, one could consider a 20 Mbps stream, representing a higher-resolution compressed video bitstream, and a 10 Mbps representing a relatively low-resolution video from the same codec operating in Constant BitRate (CBR) mode. Assuming an Ethernet frame size of 1500 B, one of the modal packet sizes for Internet transmission, with IP/UDP headers. Then, the former bitstream generates about 1698 packets/s, while the figure for the latter bitstream is 849 packets/s. In terms of encoded data, the proportion of video data contributing to the same-sized display (though with different resolutions) is less per packet for the 20 Mbps stream than it is for the 10 Mbps stream. Hence, for equivalent percentage packet losses, assuming sufficient bandwidth, a greater number of higher resolution packets survive than those from the lower resolution stream. There are many other contingencies that will affect the outcome, but as a reason for investigating what the actual response is according to resolution the case is clear.

Notice that in the previous brief discussion, the assumption was for CBR transmission. CBR video (or piecewise CBR in statistical multiplexing) is more convenient for reserving bandwidth and storage space but can be less convenient in a test scenario comparing video qualities, as in CBR the underlying coding quality can vary to match the bitrate. Therefore, we have performed two sets of experiments in this paper, which apart from differences in modeling and codec configuration, also differ between VBR and CBR. As the paper brings out there are differences in interpretation of the results, depending on whether Variable BitRate (VBR) or CBR is employed. VBR video open-loop encoding is more efficient in terms of bandwidth and storage requirements than CBR closed-loop encoding. However, when streaming VBR presents more of a challenge because of variations in the dynamic range of the bitrate. On the other hand, CBR may result in poor quality when complex scenes occur because the QP is increased (resulting in lower quality) to enable the bitrate restrictions to be met. In our tests, high motion occurs for some of the test video sequences, which can result in an encoder in CBR allocating a higher QP. (Scene changes, which also can result in lower quality CBR, do not generally occur in test sequences.) In [24], it is also demonstrated that temporal variations in video quality are more apparent in HD video. Therefore, in this paper experiments are presented both for VBR and CBR video.

The remainder of this paper is organized as follows. Section 2 mentions related work in addition to that in this introductory Section or describes already mentioned work in a little more depth. Section 3 describes the methodology, video configurations, and other aspects of the video quality evaluations in the two sets of experiments reported on in this paper. Section 4 presents the video evaluations for the two sets of experiments and in doing so analyzes the results. Finally, Section 5 picks out some key points arising from the evaluations and makes recommendations for further research.

2. Related Work

As noted in [25], there has been a surge in research into video streaming Quality-of-Experience (QoE) [26]. Though the research of [25] did employ a simplified form of real-time subjective testing, it concentrated on the impact of delay and jitter. The main insight was that inter-domain jitter, which was related to periods of end-to-end packet delay, contributed to poor QoE when it occurred. From the point-of-view of choosing between the HAS and IP/UDP/RTP varieties of streaming mentioned in the previous section, HAS's use of TCP appears to imply more delay for Internet video streaming, unless peer-to-peer streaming is employed. However, there is also the issue of PLV which in [27]'s subjective tests of H.264 video was found to be dependent on: the initial mean square error between the error-free and concealed MBs in a slice; and the maximum number of partitions of a block; and the frame-type, though other factors were also significant. In view of comments in section 1 on user reaction to video quality, it was found by the same group working with MPEG-2 video [28] that even small distortions in quality as a result of packet loss resulted in user dissatisfaction. Comparing [28] with [27], because motion-compensated error concealment was employed before evaluation in the latter, the amount of motion present in lost packets was not such an important factor. All the same, from this paper's standpoint, [27]'s H.264/AVC work was conducted with low-resolution imagery, namely Source Input Format (SIF) resolution (352×240 pixels/frame). In addition, Nightingale *et al* [17], analyzed network impairment on HEVC encoded video streams below HD resolution, while [18]'s study did not extend beyond Common Intermediate Format (CIF), possibly owing to limitations imposed by target application.

Other works, as well as [27] [28], have tried to identify factors that impact upon subjective video quality. For lower resolution video, [29] identified packet loss and bitrate to be more important than frame rate. The research reported in [30] broadly examined two factors: the data loss pattern; and the content characteristics. However, from the standpoint of resolution, the subjectively tested video sequences were confined to SD (at 25 fps). Packet losses were mainly in I-frames, which reduces the generality of the tests. When scene changes occurred, causing an additional I-frame to be inserted within a Group-of-Pictures (GoP), provided the error burst was not long enough to affect both frames, scene changes actually halted the usual temporal error propagation. That finding was similar to the results of [31], which also noted that the presence of camera motion (zoom and pan) increased the subsequent distortion resulting from packet loss. Otherwise, the extent of I-frame packet loss visibility depended on the burst length in [30]. The impact of data loss was found to be dependent on the data loss pattern, especially the number of packets affected, which is consistent with [32].

Spatial resolution does not feature strongly in the evaluation criteria for some studies. For example, as mentioned in section 1 in [18] some of the set of test videos had a resolution as low as Quarter Common Intermediate Format (QCIF) (176×144 pixels/frame) with few packets per frame. That research considered wireless network errors and video content dependency and their impact on video quality. In order of importance, PLV was found to depend on content type, sender bitrate, block error rate, and mean error burst length. Though the authors of [18] acknowledge that spatial resolution does have an impact on overall quality, even for low-bitrate video, as shown in [33], they did not choose to incorporate resolution into their video quality prediction model. Thus, because inappropriate resolutions were employed, the experimental configuration appears not to be relevant to future video streaming, even that directed at mobile devices, which appear to be following the trend towards higher resolutions, already apparent for desk-top machines. In addition, in practical implementation of the video-quality prediction model [18], Peak Signal-to-Noise Ratio (PSNR) was reverted to. Though PSNR is convenient to calculate, it is not directly related to human perception [34]. For example, PSNR does not account for the masking of distortion by the presence of texture, which SSIM does expose. However, in [35], PSNR was again the basis of video quality assessment for video streaming over a lossy wireless channel. The authors of [35], specify a method of apply PSNR across a video sequence with missing frames owing to packet dropping. Further by taking into account other network imposed impairments such as the distorted frame rate an adjusted version of PSNR was arrived at. The result was found to be a metric with 0.9 correlation with MOS scores in tests.

The following studies indicate interesting future directions for video quality assessment, though they are not primarily concerned with PLV. Assessment of synthetic 3D images was addressed in [36]. The authors of [36] introduced a method for including the temporal flicker that arises from depth image distortion. The method's value was confirmed by extensive subjective tests. The database of videos in which temporal flicker is a significant issue will be most helpful in future assessments of PLV in 3D video. Video quality based on knowledge of the brain's processing of video was introduced in

[37]. The response to additive impairments and detail losses was assessed separately. Then the images were compensated for motion contrast sensitivity and visual masking before assessment. The assessment also modeled expected eye movement. A weighted metric was then calculated to achieve good predictive performance. The work in [38] is particularly innovative in that the brain's electrical activity was compared with the conscious response of video viewers. The study found that the brain appeared to detect some impairments to the video that the viewers did not consciously respond to. The implication is that subjective testing in the future may be conducted by direct measurement of brain activity alone.

3. Evaluation Methodology

To judge the relationship between PLV and spatial resolution, two sets of experiments were performed. The first set concentrated on modeling packet loss by simulation means, while the second set performed a live streaming experiment to judge the effect of packet scheduling when packet losses occurred. In practical terms, the authors originally performed a set of simulations, which they then sought to extend and endorse by including 3D video in a live streaming scenario, along with a change in streaming mode. It is never possible to examine every configuration of wireless video streaming but in the two sets of experiments there is scope for extrapolation by the reader to a configuration of interest in order to judge the likely impact of changing spatial resolutions. The test conditions for the two sets of experiments are summarized in Table 1.

In both sets of experiments, the content characteristics from the encoding perspective were found through the spatio-temporal classification metrics in recommendation ITU-T P.910 [40]. This classifier establishes a spatial index (SI) from a video sequence by taking the luminance magnitude of a Sobel filter's output and forms a temporal index (TI), based upon successive frame differences using the luminance values. The advantage of this method of classification is that it can be performed in real-time, possibly using the software tool mentioned in [41]. TI can range from 0 to 80, with 0 meaning very limited motion and SI can vary from 0 to 250, with 0 implying very little spatial detail. Notice that [41] also introduces a database of H.264/AVC video bitstreams with packet losses for the purpose of subjective testing. However, the maximum resolution is 4CIF (704×480 pixels/frame), making the database inappropriate for this study, as results in 4CIF cannot be compared

Table 1: Experimental test conditions

Condition	Experiment one	Experiment two
Packet loss modeling	By simulation	By live streaming
Spatial resolutions	SD, HD	SD, HD, 4kUHD
3D	SD, HD	-
Codec	H.264/AVC	H.264/AVC, HEVC
Codec(s) implementation	Joint Model(JM) v.15.1 with FRExt extensions	FFmpeg
Streaming mode	VBR	CBR
Network protocols	IP/UDP	IP/UDP
Maximum data-link packet size	1.5 kB	1.5 kB
Multimedia packaging	RTP	MPEG-2 TS [39]
Error concealment technique	Previous frame	Previous frame

with results for higher resolutions. In both sets of experiments also, previous frame error concealment was employed, avoiding introducing, if a more sophisticated form of error concealment was in place, the added factor of the form of error concealment.

3.1. Testing by simulation

For these tests, three video sequences with similar spatial complexity but increasing motion activity were selected for encoding into SD, HD (1024×720 pixels/frame, i.e. 4:3 aspect ratio), and 3D in the two resolutions. Figure 1 shows a sample frame and depth image for the 3D version of the selected video sequences, while Table 2 shows their associated content classification. Each sequence of 200 frames in length was captured in YUV 4:2:0 chroma format at 25 fps, i.e. 8 s in duration. The 3D depth sequence was encoded with the same QP in order to avoid issues arising from asymmetric assignment of the QPs.

Coding parameters are summarized in Table 3. The GoP size was 16 with one intra-coded I-frame and the remainder of the frames being predictively coded P-frames. This arrangement is commonly employed for streaming to mobile devices in order to reduce the coding complexity and memory references arising from bi-predictive B-frames. Context Adaptive Variable-Length Coding (CAVLC) entropy coding was adopted to speed-up tasting, and rate-distortion analysis was turned off for the same reason. Notice that an H.264/AVC profile, with its IDC code number given in Table 3 is a set of encoding capabilities, while coding levels are requirements upon a decoder's performance, e.g. decoding speed.

Assessing video quality for the 2D versions of the three sequences was straightforward. Each frame was evaluated with VQM software [5] before forming an arithmetic mean of the VQM scores. To allow comparison across the first and second set of experiments, the VQM rating was converted into an objective MOS scale [42]. This conversion can be accomplished by applying equation (1).

$$MOS = 5 - 4 \times VQM \tag{1}$$

where VQM and MOS are the VQM and MOS values respectively. The mapping arises because MOS has five quality grades ranging from 1 to 5, the maximum value, whereas VQM usually ranges from 0 (original quality) to 1 (severe distortion). Though in the original VQM paper [5], the maximum VQM value is taken to be 1.2, values beyond 1 rarely occur in practice, which accounts for the form of (1).

To assess 3D video quality the 2D color images were assessed just as before, i.e. by means of VQM. However,the depth-quality model of [43] was adopted in preference to the direct method of [42], which, intended for asymmetric QPs, assesses the rendered left and right images separately. Thus, depth-map assessment took advantage of the depth-quality map of [43] based on estimating the quality of the depth signal, which in turn requires identification of the principal depth planes. Subsequently, the overall 3D quality was established by the mathematical model of [43], which combines the VQM values and their corresponding depth-map value. This quality scale, denoted as Q, is on a continuous scale from 0 (complete loss) to 1 (original quality). The conversion to objective MOS follows according to (2).

$$MOS = 5 - 4 \times (1 - Q) \tag{2}$$

In detail, the procedure to find Q is as follows. A visual depth image is initially segmented into depth planes. This is achieved through a histogram of the pixel values in a depth image. In the histogram, peaks are identified, as these represent different depth planes. The peaks in the reference image are statistically compared with the peaks in the distorted image.The comparison results in three different measures [44]: M_1, which measures the distortion of the relative distance within each depth plane; M_2, which measures distortion

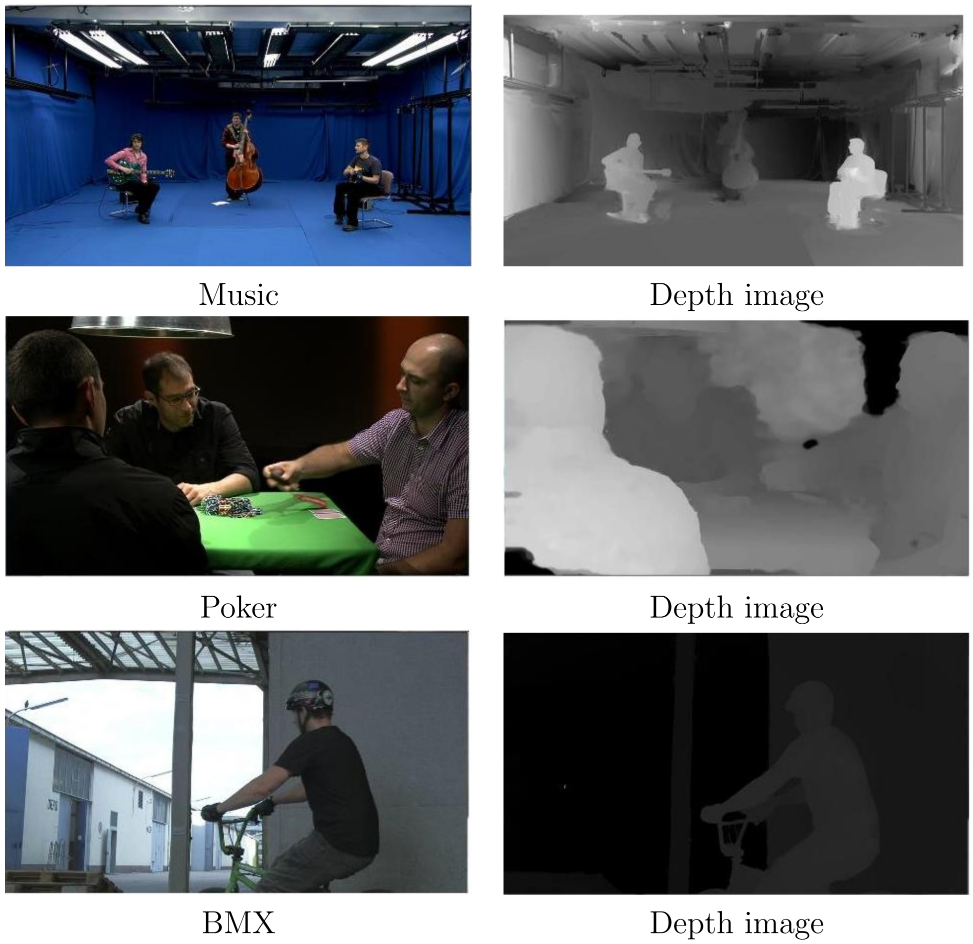


Figure 1: One representative frame from each of the three source video sequences for experiment one

Table 2: Video sequences content type for experiment one

Video sequence	SI	TI	Motion
Music	48.69	4.83	Low
Poker	53.26	12.17	Moderate
BMX	56.01	23.25	High

Table 3: H.264/AVC encoding parameters for experiment one

Parameter	Value
Profile IDC	High (100)
Level IDC	30 (SD), 32 (HD)
GoP structure	IPPP...
Group of Pictures(GOP) size	16
Entropy coding	CAVLC
Search range	32
Slice mode	Packetized (bytes)
Output file format	RTP packets
Rate control	Disabled

in the consistency of each depth plane; and M_3, which determines the structural error of the depth. These three measures are subsequently combined to form the Mean Disparity Distortion Model (MDDM) value. MDDM contributes to the overall 3D quality, Q. That contribution is governed by the activity in the depth map. To find the contribution over a video sequence, the sequence is segmented into blocks of N frames, where N is the frame rate. The standard deviation of the value of each pixel position per block is found. Then the arithmetic mean of the standard deviations is calculated to form the Z-direction Motion Activity (ZMA). ZMA is next normalized according to the resolution of the depth map and the range of the pixel luminance to give normalized ZMA (nZMA). Q is actually found through (3).

$$Q = f_1(content).ImageQuality + f_2(content).DepthQuality, \qquad (3)$$

where $f_1 + f_2 = 1$. In our experiments, $ImageQuality$ was set to the VQM value and $DepthQuality$ was set to the MDDM value. Finally, f_1 and f_2 were formed using the formulas found in [43] as a result of subjective testing, namely (4) and (5) respectively.

$$f_1 = 1 - 0.997nZMA^{0.2393} \qquad (4)$$

$$f_2 = 0.997nZMA^{0.2393} \qquad (5)$$

Burst packet losses were introduced into the simulated packet stream, according to the Gilbert-Elliott model (a simplified two-state hidden Markov model). The hidden state was modeled by a Uniform distribution in order to

Table 4: Simulated parameters for experiment one

Parameter	Value
Content type (CT)	Low, moderate, high TI
Spatial resolution (R)	SD (720×576)
	HD (1280×720)
Encoder QP	16, 24, 32
Packet Loss Rate (PLR)	0%, 1%, 2.5%, 5%, 7.5%
Mean Burst Length (MBL)	1 to 7

produce various Mean Burst Lengths (MBL)s. The intention of this common procedure is not to emulate a physical wireless channel, as (say) sampling from a Rayleigh distribution might do, but to accurately reproduce [45] an application receiver's experience of packet loss from fading.

The resulting simulation regime is summarized in Table 4. In that Table, an MBL of 1 signifies random packet losses, while as the MBL rises to seven so do the 'bursty' conditions. The QP values recorded represent high, medium, and low quality from an H.264/AVC range of 0 to 51. Each of the conditions in Table 4, i.e. CT, R, QP, and PLR, were each simulated ten times by starting from a different position in the video trace file on each occasion. Standard deviations of the results were taken to check the validity of the results, with for the most part error bounds plus or minus a few percentage points of the MOS values.

3.2. Testing by Transmission

For these experiments, the choice of source test sequence was constrained by the availability of 4k versions. Thus, in some cases 4k 'adaptations' of well-known test sequences were employed well-known, reference video sequences, i.e. *Coast*, *Foreman*, and *News* were selected. Figure 2 shows representative frames from the 4K versions. Additionally, *Sintel 4k* was accessed at [46]. In Table 5, these are arranged in order of their motion activity, similarly to Table 2.

In terms of choice of transmission parameters, choice of the lowest CBR bitrate for 4kUHD video was guided by the approximate savings of 35.4% of an HEVC codec over an H.264/AVC codec reported in [19]. As mentioned in Section 1, adjusting for coding gain between the two codecs is in line with Pinson et al's [6] previous comparison of MPEG-2 and H.264/AVC equal

Sintel | Coast

News | Foreman

Figure 2: One representative frame from each of the four source video sequences for experiment two

Table 5: Video sequences content type for experiment two

Video sequence	SI	TI	Motion
Coast	10.84	16.92	Moderate
News	17.52	21.24	Moderate
Foreman	19.71	38.29	High
Sintel	16.39	72.26	High

Table 6: 4kUHD parameters for experiment two

Parameter	Value
Packet Loss Rate (PLR)	0.1 %
PLR (0.1%) + 4kUHD HEVC bitrate	13.5 Mbps, 18 Mbps, 23 Mbps, 25 Mbps

expected video quality. Further, in-house results [47] indicated that 4kUHD transmission is possible over recent wireless links (IEEE 802.11n [48] and IEEE 802.11ad), because H.264/AVC can compress to 20 Mbps if average bitrate rate control is utilized. Thus, a rate of around 13.5 Mbps for 4kUHD with an HEVC codec is arrived at, when an additional anticipated 0.1% PLR is anticipated. Consequently, Table 6 reports the transmission parameters for 4kUHD video transmission over a Wide Area Network (WAN). An SCE WAN emulator [49] using bridging mode was attached to the outbound link of the sender, which added MPEG-2 TS/UDP/IP headers to the bitstream payload. The tests data points were all the mean of ten tests in an isolated network.

Notice that there are normally seven MPEG-2 TS packets per one UDP transport layer packet dropped by the WAN emulator. Selected codec configurations are shown in Table 7 for CBR streaming. It will be seen that larger GoPs were employed in the second set of experiments and the GoP frame structure was also varied. One principal way HEVC differs from H.264/AVC is in employing Coding Tree Blocks (CTBs) in a quadtree structure in order to extract greater coding efficiency when coding homogeneous areas of a video frame. On the other hand, in the FFmpeg implementation of both H.264/AVC and HEVC, the same rate control method is used. This is the fast rate-control method described in [50], originally for the x264 implementation of an H.264 standard codec.

4. Experimental results

4.1. Findings from experiment one

To judge the impact of packet loss, encoding loss with zero PLR was found for SD and HD resolutions in 2D and 3D formats, refer to Figure 3. For the 3D assessments, the Q metric was converted to the MOS range (1 to 5) by (2). The effect of increasing the QP across all three video clips is to reduce the quality, as expected. When including the depth images in the 3D evaluations,

Table 7: Codec parameters for experiment two

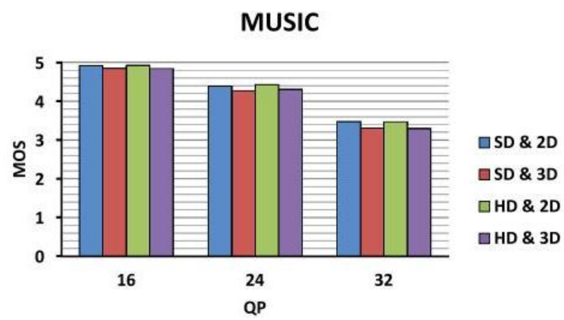
Parameter	HEVC	H.264/AVC
Profile	Main	High (5.1)
Processing unit	Coding tree block	Macroblock
Processing unit size	64×64	16×16
GoP size	25	40
GoP frame structure	IBBPBBP...	IPPP...
CBR bitrates	13.5 Mbps	20 Mbps

the overall quality is diluted compared to its 2D equivalent. Referring back to Table 2, it appears that motion-activity is not a guide to quality even when using VQM. (SSIM does not include any motion assessment, as it was originally designed for still images and simply takes the mean evaluation values across a video's sequence.) For example, Music at QP=32 is of lower assessed quality than high motion BMX.

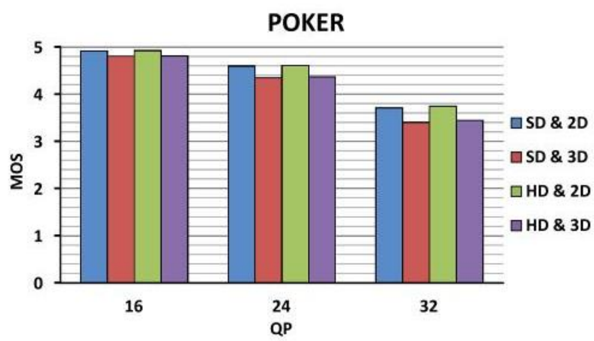
4.1.1. Results with isolated packet loss

For 2D imagery it was found, compared to no packet loss assessments, that the differences between the QPs was smoothed out by packet losses, even when these occurred in isolation. This is evident from Figure 4 across the range of PLRs tested. The implication of this finding is that decreasing the QP by a small amount so as to improve the delivered quality may be difficult to justify, as packet losses can remove any gains, while an increased bandwidth will also result. Conversely, for non-real-time or one-way streaming it may be better to trade quality for content protection by means of a reliable underlying protocol, usually TCP. If packet losses do occur then HD appears able to tolerate packet losses better than SD does. That finding confirms the intuitive argument in section 1 about the relative reduction in impact from losing a packet in HD. The objective MOS ratings do not take into account differences in viewing conditions, such as viewing difference, between SD and HD.

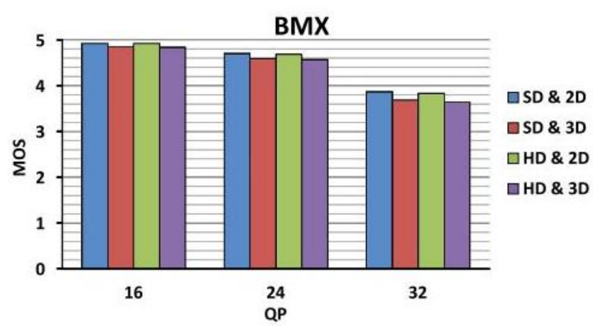
3D assessments in Figure 5 followed a similar trend to the no loss evaluations, i.e. quality was reduced by including a depth image but the response pattern for random losses was similar to the 2D findings.



a)

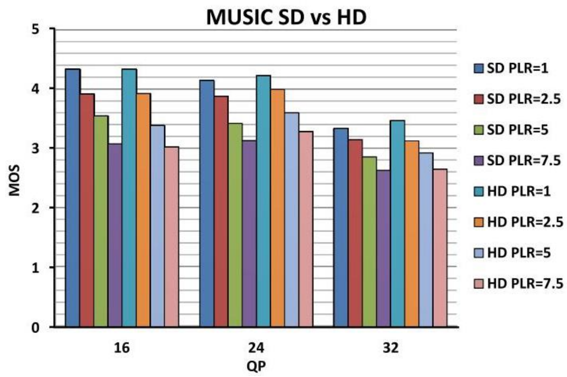


b)

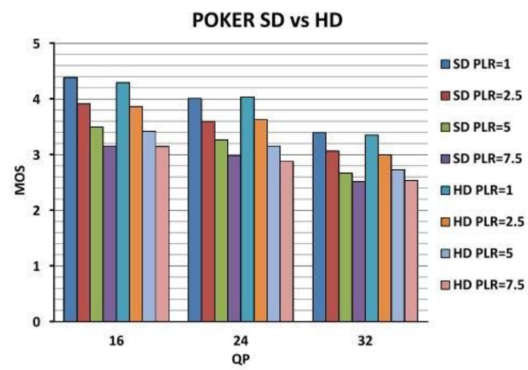


c)

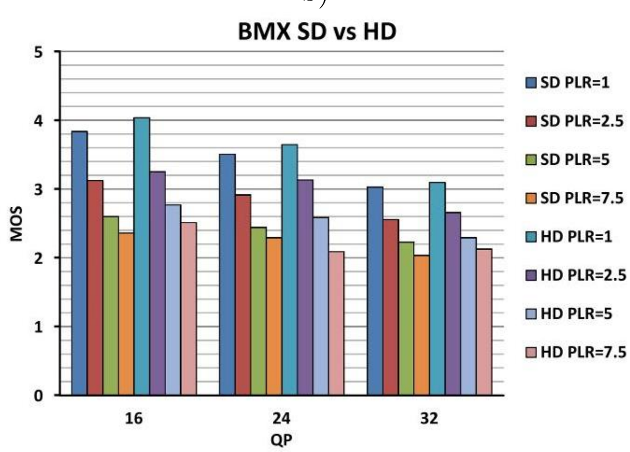
Figure 3: Objective MOS video quality assessment for SD, HD, and 3D sequences with PLR=0, a) Music, b) Poker, c) BMX



a)

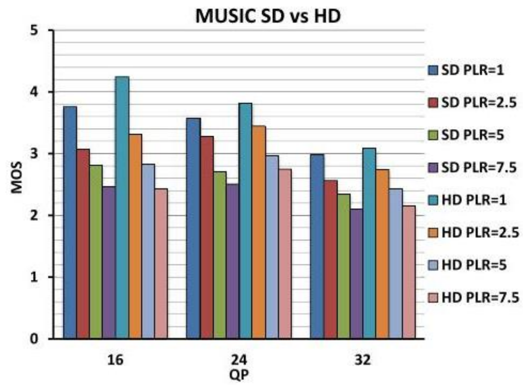


b)

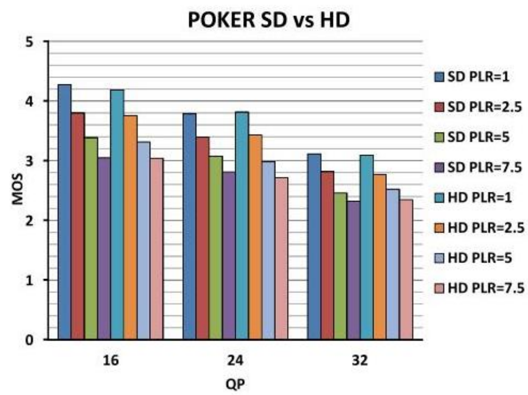


c)

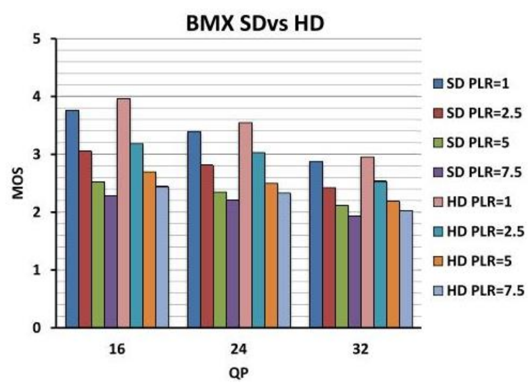
Figure 4: Objective MOS video quality assessment for SD and HD with varying PLR, a) Music, b) Poker, c) BMX



a)

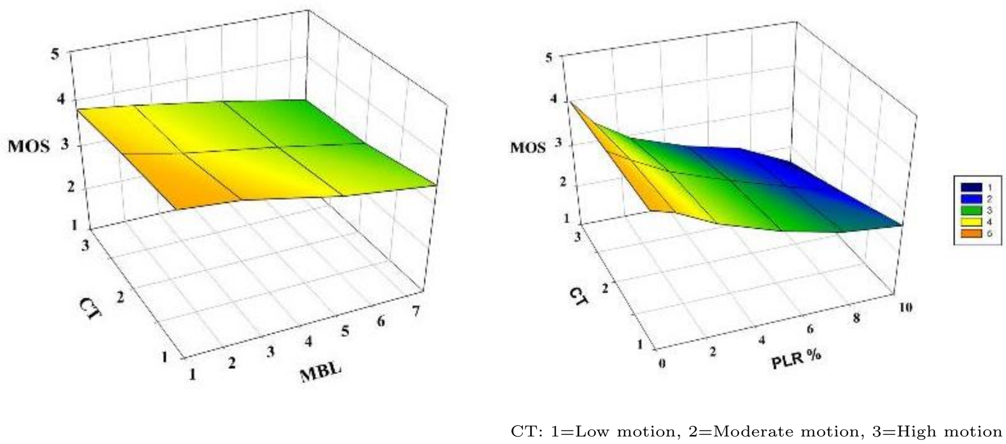


b)



c)

Figure 5: Objective MOS video quality assessment for 3D SD and HD with varying PLR, a) Music, b) Poker, c) BMX



a) b)

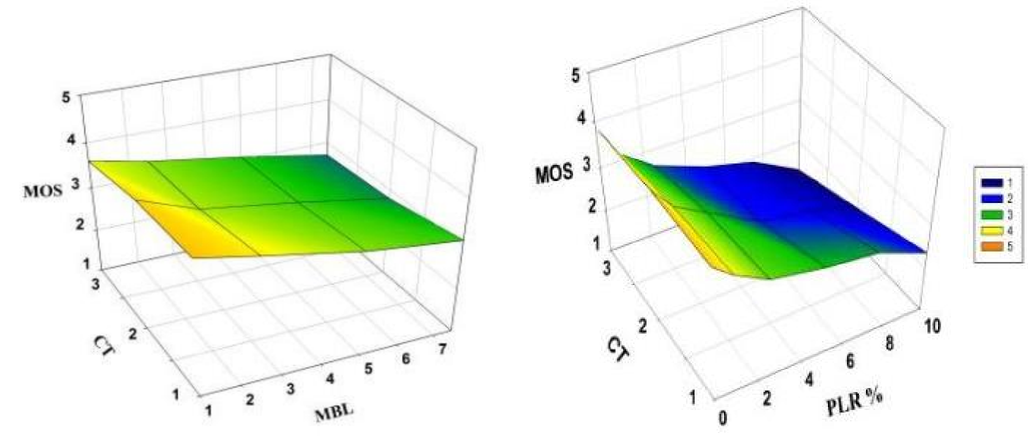
Figure 6: Influence on 2D CT of a) varying MBL for fixed PLR = 0.1%, b) varying PLR for MBL = 1. The key defines the color mapping to objective MOS scores

4.1.2. Results including 'bursty' packet loss

The effect of CT was compared with either keeping the the PLR fixed (at 0.1%) and increasing the burst length or retaining isolated packet losses but increasing the PLR. These findings are summarized for 2D and 3D HD video clips in Figs. 6 and 7 . In this case, it was evident that quality was very dependent on the CT. For low motion video quality was acceptable up to a PLR of 4%. For high motion video, the findings indicate that once a threshold of 2% PLR is reached, the quality starts to become unacceptable. Compared to the impact of PLR, MBL has a less clear effect (albeit at the low PLR tested). It appears that as the packet burst length increases, their influence upon inter-frame dependencies decreases, due to a relative reduction of the impact upon temporal error propagation.

4.2. Findings from experiment two

Again in order to calibrate later findings, tests were conducted to evaluate the video quality without packet loss impairment. In these tests, reported in Fig 8 for HEVC CBR at 13.5 Mbps and H.264/AVC at 20 Mbps, the findings suggest that HEVC produces marginally (as far as the viewer is concerned) better quality video at a lower bitrate. These results appear to differ from experiment one's results in two ways: 1) higher resolution video results in lower objective MOS ratings; and 2) lower motion videos



CT: 1=Low motion, 2=Moderate motion, 3=High motion

a) b)

Figure 7: Influence on 3D CT of a) varying MBL for fixed PLR = 0.1%, b) varying PLR for MBL = 1. The key defines the color mapping to objective MOS scores

have higher qualities. However, the results are not contradictory, because, in this set of tests, the CBR was fixed whatever the resolution. This naturally results in less compression for lower resolutions, as the QP varies to match the available bitrate. Moreover, the encoder's have been able to take advantage of the additional bitrate to improve the quality of low-motion videos. In other words, both codecs avoid simply increasing the bitrate artificially, by, for example, including more intra-coded CTBs or MBs. (Notice that adding intra-coded blocks usually takes place when high motion reveals areas in the reference frame which cannot easily be matched through inter-coding, whereas here we have low motion.)

Figure 9 reports random packet loss of 0.1% for a selection of the video sequences. The H.264/AVC codec now (compared to Figure 8) appears more resilient to packet loss than HEVC resulting in higher quality ratings for H.264/AVC encoding. This finding strongly suggests that HEVC's more efficient encoding makes it more sensitive to packet loss, once the relative coding gains have been allowed for by scaling the CBRs (with HEVC at 13.5 Mbps and H.264/AVC at 20 Mbps). Packet loss now serves to exaggerate the difference in quality already starting to show in Figure 8 between lower (higher quality) and higher resolution (lower quality). Similarly, when packet losses occur, even at a low rate of 0.1%, higher motion sequences

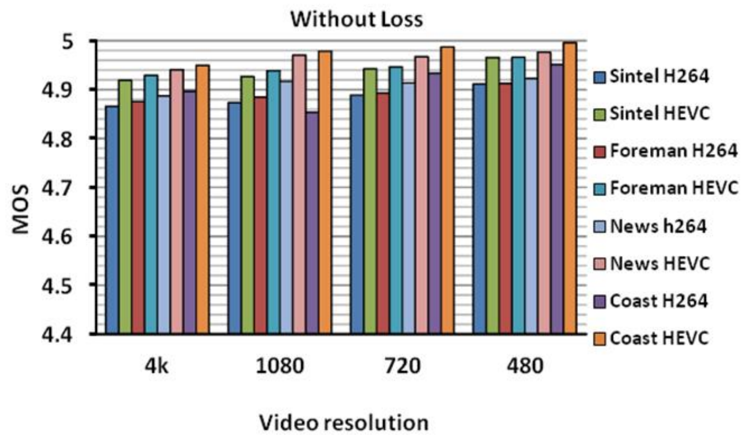


Figure 8: Objective MOS video quality assessment for a range of resolutions with PLR = 0 and either H.264/AVC or HEVC codec

such as Sintel suffer in quality much more than lower motion sequences such as Coast. Again a possible explanation, when a fixed CBR is involved, is that lower spatial resolution and lower motion video sequences will tend to be less compressed. Consequently, with more coded content per packet for lower resolution and lower motion videos, error concealment is better able to reconstruct missing packets. In addition, at higher QPs (lower quality), packet loss has more of an impact as the effect of temporal error propagation and difficulty in performing error concealments are compounded. Lastly, the relationship between packet loss and bitrate is demonstrated in Figure 10 for 4kUHD resolution video. As the bitrate is increased, the trend is for the packet loss impact to reduce so that the video quality increases. This effect can be attributed to the amount of coded information distributed amongst the packets. For example, a 13.5 Mbps video stream will have fewer packets over time that a stream at a higher rate but this means that the amount of coded information per packet is higher compared to (say) a 25 Mbps stream. Thus, the amount of coded data in a single packet could be split into two or more packets at a higher rate. Consequently, the trend is that the sensitivity to packet loss is reduced.

4.3. Packet reordering and duplication

This section investigates the effect of packet reordering or packet duplication on 4kUHD resolution video quality encoded with HEVC, as this

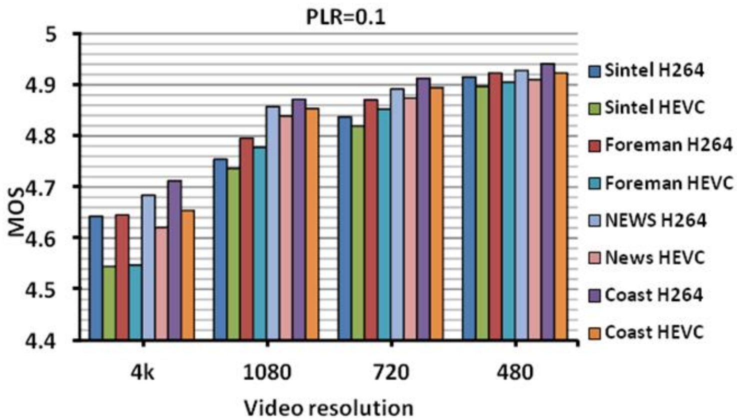


Figure 9: Objective MOS video quality assessment for a range of resolutions with PLR = 0.1% and either H.264/AVC or HEVC codec

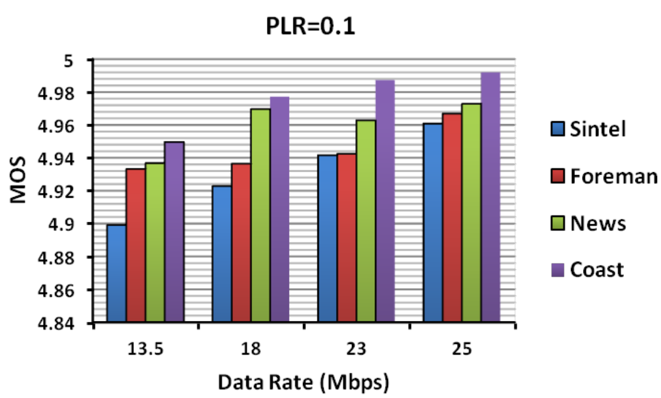
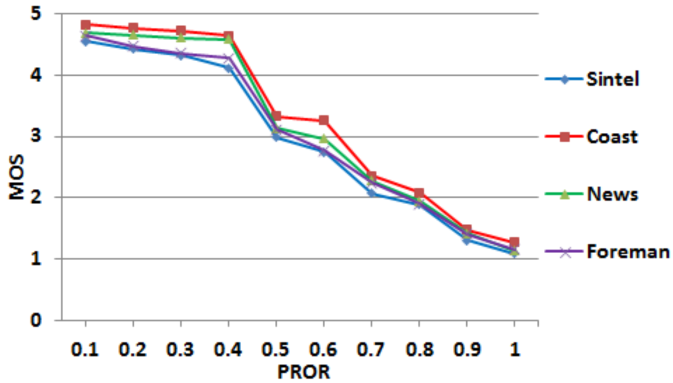


Figure 10: Objective MOS video quality assessment for a range of datarates with PLR = 0.1%, 4kUHD resolution, and encoding with the HEVC codec. Notice the truncated vertical scale.

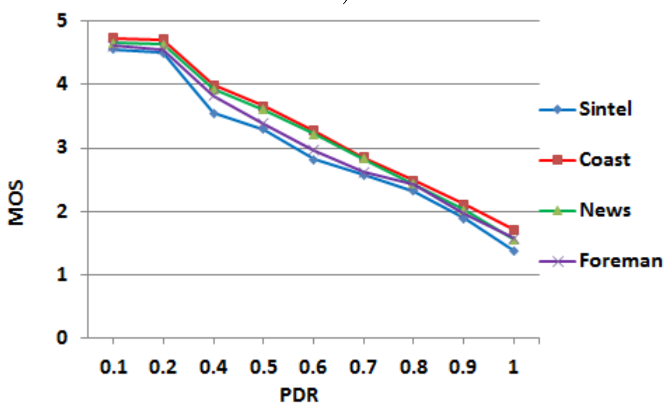
can cause additional quality degradation, the potential for which the reader should be aware of. According to [51], packet reordering in the Internet is common and can be between 3 to 5% [52]. The causes of packet reordering include load balancing, multiple network paths, and dynamic route generation [53]. The main cause of packet duplication is faults in switching/routing equipment within the network [54]. The duplication rates are usually small, e.g. 66 packets in 20,000 flows, though [54] mentions measurements of a rate over one route of 10%, owing to an incorrectly configured bridge device.

In the tests, the HEVC codec configuration was the same as that in Table 7, while the transmission and receiver/decoder set-ups were the same as for experiment two's packet loss tests. The FFmpeg decoder buffer size was set to the default of 28 KB. The SCE WAN emulator allowed random reordering/duplication rates from 0.1% to 1.0% to be set, as a point of comparison with the impact of random packet losses at the same rates. Again, there were seven MPEG-2 TS packets to each UDP packet. The main impact of packet reordering is upon delay in arriving at the decoder, which can result in dropped packets if any arrival times exceed display deadlines. As duplicate MPEG-2 packets bear the same timestamp, delay at the decoder arises because of the need to drop the duplicate packets. The 4kUHD resolution will clearly have an impact on the processing time at the decoder. This can result in later packets missing display deadlines.

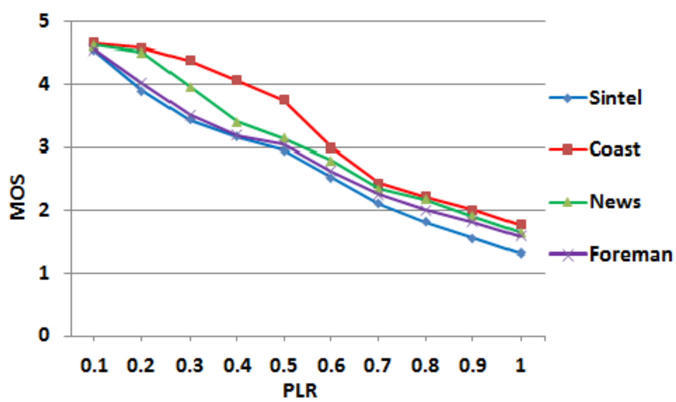
Figure 11 shows the findings for 4kUHD video. It can be observed that the impact of the PROR is less than that of packet loss, Figure 11c, at the lower reordering rates other than 0.1, though there is a content dependency effect. Figure 11b for PRORs implies that for less active video sequences, such as Coast and News (refer to the T1 column in Table 5), can accommodate a maximum 0.4% PROR before the quality drops steeply. (There may also be a threshold effect related to buffering sizes.) However, the more active two sequences experience more loss of quality after a rate of 0.1%. Thus, for safe streaming the expected PROR would need to be no higher than 0.1%, which may be difficult to accomplish, given reordering rates reported in [51], unless (say) large buffers are in place to mitigate the potential effect. However, larger buffers result in longer start-up/pre-roll delay unless adaptive buffer management [55] is employed. Plots in Figure 11c suggest that all sequences can tolerate at least up to 0.2% PDR, before the video quality deteriorates steeply beyond the rate of 0.4%. However, as previously mentioned, duplication rates beyond 0.2% are rare if not unknown. However, this analysis underestimates the potential effect upon the quality of 4kUHD



a)



b)



c)

Figure 11: Objective MOS video quality assessment for a range of rates with HEVC codec encoding and 4kUHD resolution a) PROR, b) PDR, c) PLR. N.B. The WAN emulator did not allow a PDR of 0.3 to be set.

video where either packet reordering or packet duplication or both occur and are combined with packet loss. The combination of all three effects can occur but further investigation is beyond the scope of the current paper.

4.4. Statistical significance of the results

An analysis of variance (ANOVA) test [56] was conducted, as such a test is useful for establishing the interactions between two or more independent variables. Table 8 shows the results obtained from an ANOVA test of our 2D video datasets, where the degrees of freedom are shown in the first column, the second column is the F-statistic, and the third column is the p-value. (The parameter abbreviations were introduced in Table 4.) The p-value is determined from the cumulative distribution function (cdf) of F [57]. A small p-value (e.g. $p = 0.01$) indicates that the video QoE is significantly affected by the corresponding parameter. A higher F-statistic corresponds to a higher proportion of the variance being caused by the independent variable(s) [56].

It can be observed from the magnitudes of the p-values that all five parameters had significant effects on the video QoE. In particular, it can be seen that the PLR had the highest influence on the QoE (p-value=0), followed by QP and CT, while the MBL had the smallest influence on the QoE. Moreover, there were interactions between each pair of parameters, each of which was significant. The two-way interactions between PLR and CT, and PLR and QP had the highest influence on the QoE. In addition, the ANOVA results showed that the combined impact of MBL and CT was also significant.

5. Conclusion

This paper has highlighted a number of considerations when video streaming at higher resolutions. When VBR encoding is employed, it was found that packet loss had less of an impact upon HD video than on SD video. Whether isolated or bursty packet loss occurred, it was the level of motion activity that was the most significant factor. In fact, whatever the resolution, if the packet loss rate is above 2% and the temporal complexity is high then there is a need for channel coding and/or error resiliency. For more static content then PLRs approaching 4% might be tolerated. As the HEVC codec has reduced its support for error resilience tools, compared to the H.264/AVC standard, channel coding becomes more important. The main alternative to protecting content is to choose HTTP Adaptive Streaming (HAS) in which no

Table 8: Five-way ANOVA for the QoE of 2D video

Source	Degrees of freedom	F-statistic	p-value
CT	2	132.724	0.0
R	3	95.354	0.03132
QP	4	159.584	0.0
PLR	5	402.172	0.0
MBL	2	65.991	0.01068
PLR+CT	9	20.218	0.1180
PLR+R	9	5.182	0.3132
PLR+QP	15	26.955	0.11068
PLR+MBL	20	30.466	0.2301
MBL+CT	5	2.868	0.1541
MBL+R	5	7.940	0.48223
MBL+QP	11	13.533	0.28568

errors can occur, because lost packets are retransmitted. However, for two-way streaming or for handling HAS on a portable device, there are issues of jitter, end-to-end delay, and implementation complexity. CBR encoding allowed a comparison between an H.264/AVC and HEVC codec to be conducted. This was achieved by normalizing the bitrates in line with the known difference in efficiencies between the codecs. However, it should be recalled in interpreting those results that HEVC streamed at 13.5 Mbps, whereas the equivalent bitrate chosen for H.264/AVC was 20 Mbps. If bitrate/bandwidth is not a restriction, such as over optical fiber or emerging high-bandwidth WLANs, then the comparisons are interesting because H.264/AVC coding is more tolerant to packet loss by virtue of less efficient coding. A content-dependent aspect, that is relative motion activity, was also observed in the CBR experiments, again making the case for careful consideration of the impact of motion activity, for example in sports video, upon the likely video quality. In terms of spatial resolutions, CBR 'appears' to favor lower resolution video, which appeared more tolerant to packet loss. However, this was because in the experiments the same CBR was used whether transmitting lower or higher resolution video, including 4kUHD video. Thus, the QP can be lower for lower resolution video than higher resolution video at the same CBR. Losing a packet from a lower resolution stream in those circumstances, has less of an impact, than losing a packet from a higher resolution stream. This highlights the need for careful interpretation of video quality results depending on the streaming modalities. A statistical analysis has confirmed the influence of all five parameters investigated, with PLR's influence highlighted. Future work should consider choice of packet size as how many packets and what their coding contribution is determines differences in the display quality. GoP structure and choice of QP are also important determinants of the eventual video quality of a video stream as a whole, once it enters a network or is transmitted. 4kUHD is a clear candidate for investigation and research in the medium to long term. Other areas of future study are more advanced scenarios, such as multi-view 3D videos for virtual reality and more advanced packet loss models in Long Term Evolution (LTE)-A (i.e., 4.5G) wireless networks with carrier aggregation. Finally, one should notice that channel coding is frequently employed to ameliorate the impact of packet losses. However, there are many forms of channel coding each of which introduces different bitrate overheads and latencies. Latency is a particularly issue for interactive applications such as mobile video conferencing. Therefore, this issue deserves further study after this initial study.

References

[1] J. Shin, P. C. Cosman, Classification of MPEG-2 Transport Stream packet loss visibility, in: IEEE Int. Conf. Acoustics Speech and Signal Processing, 2010, pp. 910–913.

[2] S. Argyropoulos, A. Raake, M.-N. Garcia, P. List, No-reference bit stream model for video quality assessment of H.264/AVC video based on packet loss visibility, in: IEEE Int. Conf. on Acoustics, Speech and Signal Processing, 2011, pp. 1169–1172.

[3] Subjective video quality assessment methods for multimedia applications, recommendation ITU-T P.910 (2010).

[4] A. Moorthy, K. Seshadrinathan, R. Soundarajan, A. Bovik, Wireless video quality assessment: A study of subjective scores and subjective algorithms, IEEE Trans. Circ. Syst. Video Technol. 20 (4) (2010) 513–516.

[5] M. Pinson, S. Wolf, A new standardized method for objectively measuring video quality, IEEE Trans. Broadcast. 50 (3) (2004) 312–322.

[6] M. H. Pinson, S. Wolf, G. Cermak, HDTV subjective quality of H.264 vs. MPEG-2, with and without packet loss, IEEE Trans. Broadcast 56 (1) (2010) 86–91.

[7] ITU-T, Bt.2020 : Parameter values for ultra-high definition television systems for production and international programme exchange (2014).

[8] M. Mehendale, S. Das, M. Sharma, M. e. a. Mody, A true multi-standard, programmable, low-power, full HD video-codec engine for smartphone HD SoC, in: IEEE Int. Solid State Conf., 2012, pp. 226–228.

[9] P. Merkle, Y. Wang, K. Müller, K. A. Smolic, T. Wiegand, Video plus depth format for mobile 3D services, in: IEEE 3DTV Conf., 2009, pp. 1–4.

[10] C. Poynton, Digital video and HDTV: Algorithms and interfaces, San Francisco, CA: Morgan Kaufmann, 2003.

[11] A. Vetro, A. Tourapis, K. Müller, T. Chen, 3D-TV content storage and transmission, IEEE Trans. Broadcast. 57 (2) (2011) 384–394.

[12] S.-H. Bae, J. Kim, M. Kim, S. e. a. Cho, Assessments of subjective video quality on HEVC-encoded 4K-UHD video for Beyond-HDTV broadcasting services, IEEE Trans. Broadcast 59 (2) (2013) 209–222.

[13] Z. Li, Y. Huang, G. Liu, F. Wang, Z.-L. Zhang, Y. Dai, Cloud transcoder: Bridging the format and resolution gap between Iinternet video and mobile devices, in: Int. Workshop on Network and Operating System Support for Digital Audio and Video, 2012, pp. 33–38.

[14] I. Sodagar, The MPEG-DASH standard for multimedia streaming over the Internet, IEEE Multimedia 18 (4) (2011) 62–67.

[15] B. Bing, 3D and HD broadband video networking, Norwood, MA: Artech House, 2010.

[16] F. Battisti, M. Carli, P. Paudyal, QoS to QoE mapping model for wired/wireless video communication, in: Euro Med Telco Conf., 2014, pp. 1–6.

[17] J. Nightingale, Q. Wang, C. Grecos, HEVStream: a framework for streaming and evaluation of high efficiency video coding (HEVC) content in loss-prone networks, IEEE Trans. Consumer Electron. 58 (2) (2012) 404–412.

[18] A. Khan, L. Sun, E. Ifeachor, QoE prediction model and its application in video quality adaptation over UMTS networks, IEEE Trans. Multimed. 14 (2) (2012) 431–442.

[19] J. Ohm, G. Sullivan, H. Schwarz, T. Tan, T. Wiegand, Comparison of the coding efficiency of video coding standards — including High Efficiency Video Coding (HEVC), IEEE Trans. Circuits Syst. Video Technol. 22 (12) (2012) 1669–1684.

[20] S. Mohamed, G. Rubino, A study of real-time packet video quality using random neural nets, IEEE Trans. Circuits Syst. Video Technol. 12 (12) (2002) 1071–1083.

[21] J. Nightingale, Q. Wang, C. Grecos, S. Goma, The impact of network impairment on quality of experience (QoE) in H.265/HEVC video streaming, IEEE Trans. Consumer Electron. 60 (2) (2014) 242–250.

[22] Final report from the video quality experts group on the validation of objective models of video quality assessment, phase ii, available: www.vqeg.org (2004).

[23] Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.

[24] B. Bing, 3D and HD Broadband video networking, Norwood, MA: Artech House, 2010.

[25] M. Ventakaraman, M. Chatterjee, Case study of Internet links: What degrades video QoE?, in: IEEE Globecom, 2010, pp. 1–10.

[26] R. Jain, Quality of experience, IEEE Multimedia 11 (1) (2004) 95–96.

[27] S. Kanamuri, S. Subramanian, P. Cosman, A. Reibman, Predicting H.264 packet loss visibility using a generalized linear model, in: IEEE Int. Conf. Image Process., 2006, pp. 2245–2248.

[28] S. Kanamuri, P. Cosman, A. Reibman, V. Vaishampayan, Modeling packet-loss visibility in MPEG-2 video, IEEE Trans. Multimedia 8 (2) (2006) 341–355.

[29] G. Cermak, Subjective video quality as a function of bitrate, frame rate, packet loss rate and codec, in: 1st Int. Workshop on Quality of Multimedia Experience, 2009, pp. 41–46.

[30] F. Boulos, B. Parrein, P. Le Callet, D. Hands, Perceptual effects of packet loss on h.264/avc encoded videos, in: 4th Int. Workshop Video Process. and Quality Metrics for Consumer Electronics, 2009, pp. 1–7.

[31] A. Reibman, D. Poole, Predicting packet-loss visibility using scene characteristics, in: 16th Int. Packet Video Workshop, 2007, pp. 308–317.

[32] Y. Liang, J. Apostolopoulos, B. Girod, Analysis of packet loss for compressed video: Does burst-length matter?, in: IEEE Int. Conf. Acoust. Speech, and Signal Process., 2003, pp. V–684–V–687.

[33] G. Zhai, J. Cai, Q. Lin, X. Yang, W. Zhang, M. Etoh, Cross-dimensional perceptual quality assessment for low bitrate videos, IEEE Trans. Multimedia 10 (3) (2008) 1316–1324.

[34] S. Winkler, P. Mohandas, The evolution of video quality measurement: From PSNR to hybrid metrics, IEEE Trans. Broadcast 54 (3) (2008) 1–9.

[35] A. Chan, K. Zeng, P. Mohapatra, S. Lee, S. Banerjee, Metrics for evaluating video streaming quality in lossy IEEE 802.11 wireless networks, in: IEEE INFOCOM, 2010, pp. 1–9.

[36] X. Liu, Y. Zhang, S. Hu, S. Kwong, C.-C. Kuo, Q. Peng, Subjective and objective video quality assessment of 3d synthesized views with texture/depth compression distortion, IEEE Trans. Image Process. 24 (12) (2015) 4848–4861.

[37] S. Li, L. Ma, K. Ngan, Full-reference video quality assessment by decoupling detail losses and additive impairments, IEEE Trans. Circuits Syst. Video Technol. 22 (7) (2012) 1100–1112.

[38] S. Scholler, S. Bosse, M. Treder, B. Blankertz, B. Curio, K.-R. Müller, T. Wiegand, Toward a direct measure of video quality perception using EEG, IEEE Trans. Image Process. 21 (5) (2010) 2619–2629.

[39] A.Adeyemi-Ejeye, HEVC MPEG TS definition in MPEG TS header (03 2014).
URL `https://github.com/FFmpeg/FFmpeg/pull/60/files`

[40] Subjective video quality assessment methods for multimedia applications, recommendation ITU-T P.910 (2008).

[41] F. De Simone, M. Naccari, M. Tagliasacchi, F. Dufaux, S. Tubaro, T. Ebrahimi, Subjective quality assessment of H.264/AVC video streaming with packet losses, EURASIP Journal on Image and Video Processing Volume 2011 (2011) 12 pages.

[42] J. JosKowicz, R. Sotelo, J. Ardao, Towards a general parametric model for perceptual video quality perception, IEEE Trans. Broadcast. 59 (4) (2013) 569–579.

[43] S. Yasakethu, S. Worrall, D. De Silva, A. Kondoz, A compound depth and image quality metric for measuring the effects of packet loss on 3D video, in: IEEE Int. Conf. Digital Signal Process., 2011, pp. 1–7.

[44] S. Yasakethu, D. De Silva, W. Fernando, A. Kondoz, Predicting sensation of depth in 3D video, Electronics Letters 46 (12) (2010) 837–839.

[45] M. Zorzi, R. Rao, L. Milstein, A Markov model for block errors on fading channels, in: IEEE Personal, Indoor, and Mobile Commun., Vol. 3, 1996, pp. 1074–1078.

[46] Blender-Foundation, Sintel 4k (2011).
URL `http://www.sintel.org/news/sintel-4k-version-available`

[47] A. Adeyemi-Ejeye, S. Walker, Ultra-high definition wireless video transmission using H.264 over 802.11n WLAN: Challenges and performance evaluation, in: 12th Int. Conf. Telecommun., 2013, pp. 109–114.

[48] E. Perahia, R. Stacey, Next generation wireless LANs: 802.11n and 802.11ac, Cambridge, UK: Cambridge University Press, 2013, 2nd edition.

[49] Softperfect connection emulator (sce), [Online] Available at: https://www.softperfect.com/products/connectionemulator/ (2014).

[50] L. Merritt, R. Vanam, Improved rate control and motion estimation for H.264 encoder, in: IEEE Int. Conf. on Image Process., Vol. V, 2007, pp. 309–312.

[51] G. Sarwar, E. Lochin, R. Boreli, Mitigating the impact of packet reordering to maximise the performance of multimedia applications, in: IEEE Int. Conf. on Commun., 2011, pp. 1–5.

[52] S. Jaiswal, G. Iannaccone, C. Diot, J. Kurose, D. Towsley, Measurement and classification of out-of-sequence packets in a tier-1 IP backbone, IEEE/ACM Trans. Netw. 15 (1) (2007) 54–66.

[53] L. Gharai, C. Perkins, C. Lehman, Packet reordering, high speed networks and transport protocol performance, in: IEEE Int. Conf. Commun. Computer Netw., 2004, pp. 73–78.

[54] C. Perkins, RTP: Audio and Video for the Internet, Addison-Wesley,Boston, MA, 2003.

[55] B. Girod, M. Kalman, Y. Liang, R. Zhang, Advances in channel-adaptive video streaming, in: IEEE Int. Conf. Image Process., 2002, pp. 9–12.

[56] G. W. Snecdecor, W. G. Cochran, Statistical Methods, Chichester, UK: Wiley, 1991.

[57] W. Krzanowski, Principles of multivariate analysis, Oxford, UK: Oxford University Press, 2000.