ADAPTIVE VBR VIDEO FOR CONGESTED WIRELESS HOME NETWORKS

Ismail A. Ali, Martin Fleury, *Member*, IEEE, and Mohammed Ghanbari, *Fellow*, IEEE
University of Essex, Colchester CO4 3SQ, UK
{iaali, fleum, ghan}@essex.ac.uk

ABSTRACT

As the era of media-player applications seems to be receding, native streaming is being reconsidered in commercial environments. Ease of deployment is important to such developers and in that respect this paper proposes quality-adaptive variable bitrate (VBR) video streaming for home wireless networks. The main intention of the scheme is to adapt to congestion, taking advantage of simple buffer monitoring and transcoded bitrates. Compared to constant quality VBR and constant bitrate streaming, the paper demonstrates advantages in terms of stable delivered quality and reduced link latency.

Index Terms— Adaptive quality streaming, CBR, VBR, congestion control

1. INTRODUCTION

In the progressive download environment of HTTP adaptive streaming [1], adaptation takes place at the video chunk-level with typical chunks being between 1 s and 10 s in duration. Unfortunately, for wireless networks, this adaptive approach has disadvantages in respect to native streaming. Each chunk has to be stored at multiple fidelities (typically six) and a complex management process is needed to select the desired chunk, as HTTP is stateless. Small chunk sizes reduce coding efficiency (for example, if each chunk forms a single Group of Pictures (GoP) with a single I-frame), whereas longer chunks are less adaptable. As TCP transport is employed, the well-known problem of retransmission delays occurs, due not from congestion but from channel drops. In a study of a Content Distribution Network (CDN) [2] it was found that, though feedback took place every 2 s, in practice the response to available bandwidth occurred gradually over a much longer time span. The application-layer Real-Time Messaging Protocol (RTMP) is normally employed with progressive download, but, being proprietary, it can be a difficult implementation.

In this paper, a form of quality adaptive streaming is also proposed but by using native streaming. Such streaming uses the Real-time Transport Protocol/Real-Time Control Protocols (RTP/RTCP) over IP/UDP, with the 'statefull' Real-Time Streaming Protocol (RTSP) for connection management across firewalls and across Network Address Translation (NAT). Employing adaptive variable bit rate (VBR) retains the advantages of constant quality and an open coding loop (in single pass coding mode) but at the same time it adapts to available bandwidth induced by wireless LAN congestion. Adaptation in this paper is achieved by network output buffer monitoring and demonstrated for streaming from an IEEE 802.11 access point (AP). As an example, a (more complex) buffer monitoring scheme for IEEE 802.15.4 (Bluetooth) networks is illustrated in [3].

Rate adaptation can either take place through bit-rate transcoding (transrating) [4] down from a high quality version or by altering the quantization parameter (QP) at a remote encoder. In fact, progressive download systems using simulcast such as that of YouTube will already use source bitrate transcoding, when the original raw YUV video is not available. In [5], a re-quantization scheme for H.264/Advanced Video Coding (AVC) requantization is proposed that replaces the H.264/AVC 6-tap filter for quarter pixel interpolation by a reduced complexity interpolation filter which achieves similar visual quality. The scheme has been demonstrated with unicast streams for ease of comparison but it is also the case that the Scalable Video Coding (SVC) extension to H.264 [6] could be used to change the rate. However, the apparent absence of a hardware implementation of SVC may be an impediment to its widespread adoption.

The review in [7] considered several modes of VBR output from an encoder: 1) unconstrained VBR (considered though not used by us); 2) smoothed video (not used for the reason given at the end of the next paragraph); 3) constrained VBR with encoder knowledge of the video buffer state and the network constraints; and 4) feedback VBR, when the encoder has knowledge of the network state, especially congestion periods. Thus, in those terms the proposed system adopts option four for streaming over wireless.

Constant Bit Rate (CBR) streaming is also possible but this can, depending on algorithm, require a closed coding loop to arrive at the desired bitrate, with problems at scene changes. CBR streaming results in quality fluctuations as a matter of course. (Such fluctuations can be more disconcerting to the user than lower quality video [8] and in our paper rate adaptations are restricted to GoP boundaries.) However, CBR is normal for conversational streaming applications to avoid delay jitter, though notice that VBR can lead to a reduction in encoder buffer size [9]. Though CBR allows prediction of storage space (and transmission bandwidth for terrestrial broadcasts) in advance it actually can consume more space than VBR [9]. Streaming VBR video without adaptation is also challenging because of the highly dynamic nature of the bitrate and, without rate smoothing, will become more of a challenge as spatial resolutions increase. Unfortunately, rate smoothing [10] may introduce additional latency, because of the need to look-ahead before adjusting the bitrate.

Of course, the observations in this paper concern packet-switched networks, as it is also possible in circuit-switched systems such as wireless ATM, to employ piecewise CBR transmission [11] for VBR-encoded video.

The contribution of this paper is to demonstrate the effectiveness of adaptive VBR for wireless communication in terms of improved objective video quality and reduced end-to-end delay. It does so in comparison with higher and low constant quality video and CBR streaming.

The remainder of this paper is organized as follows. Section 2 details the HN scenario and includes some tests to judge bitrates and computational overhead. Section 3 demonstrates the advantages of the adaptive VBR scheme within a home network (HN), while Section 4 rounds off the paper.

2. METHODOLOGY

The adaptive VBR scheme was tested in the home network [12] scenario illustrated in Fig. 1. The scenario was simulated using the well-known ns-2 network simulator. To avoid rewiring issues in the HN, IEEE 802.11 systems frequently distribute video content, which must compete with other traffic such as best-effort HTTP and background file transfers.

In the scenario of Fig. 1, stored video is distributed from a Network Attached Storage (NAS) unit connected to the wireless router. The NAS video server has cross-layer access to the 802.11b buffer occupancy level. File download with UDP or TCP transport from an external source results in CBR cross-traffic, causing congestion. The external CBR source varies its rate in piecewise fashion over time, as illustrated in Fig. 2 for the 35s of the test sequence. This CBR source is delivered to the Notebook device in Fig. 1. The widely adopted shadowing propagation model [13] was employed with the path loss exponent set to four, resulting in favorable wireless channel conditions on the 802.11 HN. Finally, the IEEE 802.11 AP output buffer size was set to 100 packets.

The *Paris* test sequence (1065 frames) was encoded at Common Intermediate Format (CIF) (352 × 288 pixel/picture), with a frame structure of IPPP….I (GoP size 15) at 30 Hz. Paris was selected for its relative longevity as a test sequence. It shows two presenters in a TV studio setting, with some motion complexity caused by movement

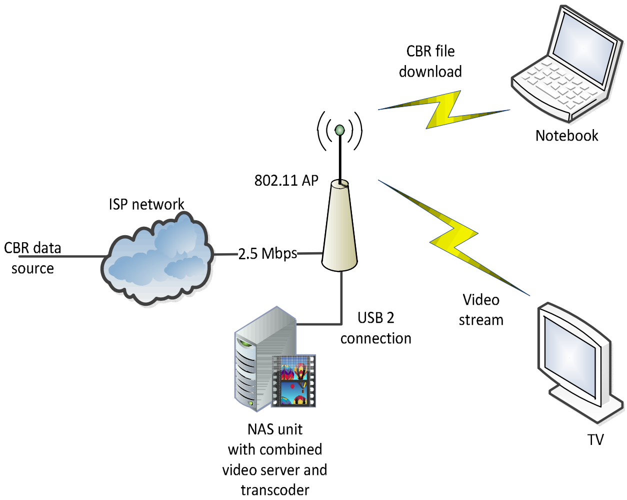

Fig. 1. Home network evaluation scenario

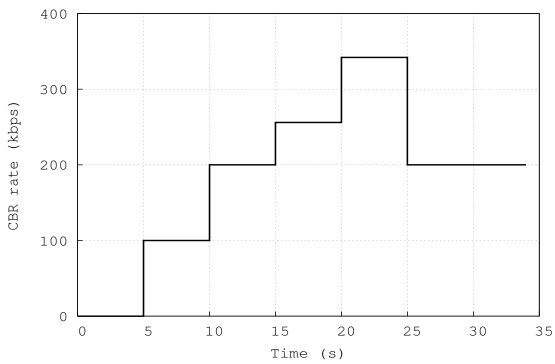

Fig. 2. Stepped CBR source acting as cross-traffic

of objects in their hands, and spatial complexity in the background. Maximum slice size was set to 1024 B.

Fig. 3 illustrates frame-by-frame input bit-rates for sample encodings of *Paris*. In Fig. 3a, the majority of the high frame sizes (approaching 20 kb) arise from the larger I-frame sizes. However, there are also two main periods when P-frames are encoded at a higher bitrate to maintain video quality when faced with increased source coding complexity. Fig. 3b shows an equivalent *Paris* CBR stream at an approximate rate corresponding to the average VBR rate of the higher of the two input unconstrained VBR rates (refer forward to Table 1). From the I-frame size fluctuations of the CBR stream in Fig. 3b, it is apparent that video quality will also fluctuate. Then Fig. 3c shows the output frame sizes of the adaptive VBR encoded sequence in response to the stepped congestion source of Fig. 2.

Before running network tests, coding wall-clock time was examined for the JM 14.2 H.264/AVC software encoder, Table 1, to compare unconstrained VBR at the highest and lowest quality when experiencing feedback with CBR video at equivalent bitrates. From Table 1, the CBR encoding time is marginally longer. CAVLC/UVLC entropy coding was employed, with Enhanced Predictive Zonal Search (EPZS) to speed up motion estimation. Both the VBR and CBR

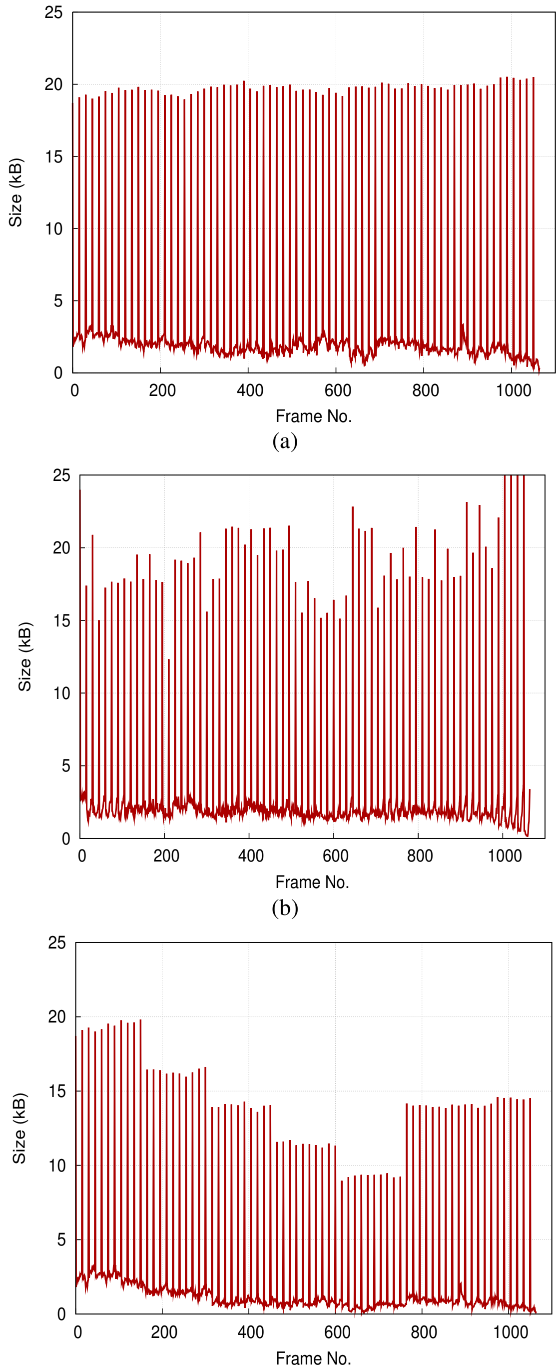

Fig. 3. Frame-by-frame compressed size for *Paris* in kbps for (a) higher-quality constrained VBR (QP=28), (b) CBR-encoded stream at 725 kbps, and (c) proposed variable quality VBR

were single-pass encoded. This sample test took place on an Intel Core 2 Quad Q6600 machine running at a nominal clock speed of 2.4 GHz. However, notice that the intention here is to give relative, indicative timings, as absolute timings are very much configuration dependent.

Table 1. Wall-clock coding times for *Paris* (1065 frames)

QP	VBR (s)	CBR (s)	Average rate (kbps)
28	397.973	411.268	725
36	365.271	365.557	260

3. RESULTS

For the HN scenario, the proposed “feedback VBR” [7] was employed with QP set at one of 28, 30, 32, 34, or 36, i.e. from higher to low quality. The proposed congestion adaptive VBR varied between those QPs. CBR was initially streamed at the average rate of the quality adaptive scheme when faced with exactly the same simulated CBR cross-traffic. Similarly, higher- and low-quality video, with the same QPs of Table 1, was streamed in independent tests across the HN. We used Peak Signal-to-Noise Ration (PSNR), which is considered [14] an adequate predictor of video quality, when comparing coding conditions for any one video clip at the same frame rate, as occurs herein.

From Fig. 4 for per-frame video quality over time, adaptive VBR streaming responds to the step changes in the cross-traffic by changing its quality, except when self-congestion causes packet drops, at the reduction in quality before the 200th frame. Notice that the adaptive VBR stream tracks the higher-quality unconstrained VBR plot up to the point when there is a sudden drop in quality. Repeated small increases in quality are clearly the result of spatially encoded I-frames occurring. As soon as the stepwise CBR cross-traffic is turned on (with UDP transport in Figs. 4 and 5), the unconstrained higher-quality VBR stream suffers severe drops in quality due to the combined impact of the cross-traffic and its own self-congestion, leading to packet drops. Low-quality unconstrained VBR never results in self-congestion but its quality is consistently lower than the adaptive VBR stream.

In Fig. 5, the highest-rate CBR stream faces similar problems to the VBR rate with QP = 28, though the onset of self-congestion occurs at a later frame. The lower quality (CBR rate 260 kbps) suffers from frequent quality fluctuations, due to the nature of the encoding mode. Though the intermediate quality CBR video (rate 475 kbps) at some frames overtakes the adaptive VBR quality, it is also prone to rapid quality fluctuations. Of course, though *Paris* is a single scene, for a sequence with scene changes, e.g. when news items were swapped between, the CBR quality fluctuations would be more severe at transition boundaries. However, notice that had CABAC entropy coding been employed [9], the CBR quality would have been somewhat improved, because the reduced bitrate from CABAC rather than CAVLC allows higher quality video for the same bitrate. This difference will become more apparent at higher bitrates, especially for high-definition streaming.

Table 2 compares mean end-to-end video frame delays[1]. One find that variable quality VBR avoids the high queuing delays of the higher-quality VBR stream, with a better overall quality than either of the other VBR streams. Mean frame end-to-end delay is higher than a CBR stream with the same mean bitrate, Table 3, but it is not prone to the disconcerting quality fluctuations arising from CBR communication, which would disrupt the HN NAS streaming experience.

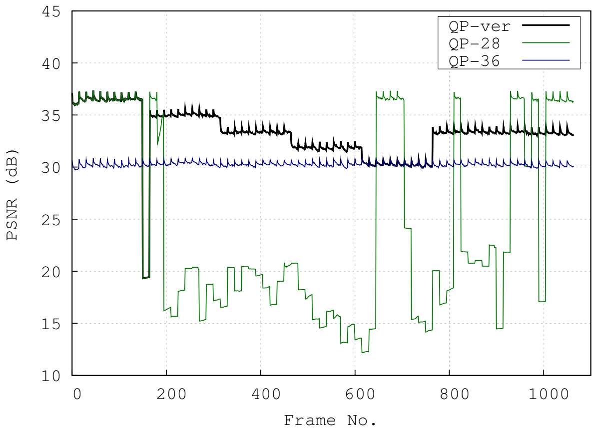

Fig. 4. VBR-encoded objective video quality (Y-PSNR) for HN video streaming at QP 28 and 36, and variable QP (QPver) with UDP congestion

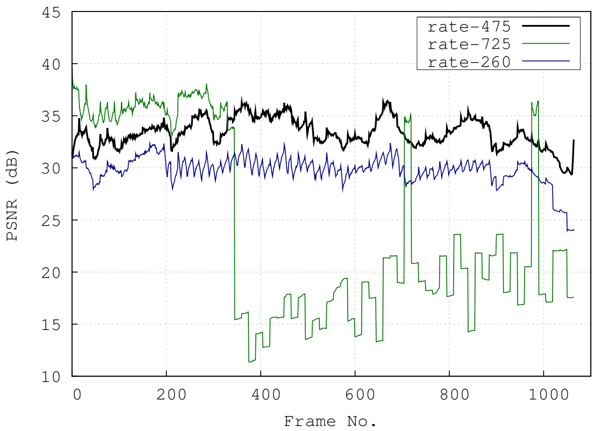

Fig. 5. CBR-encoded objective video quality (Y-PSNR) for HN video streaming at rates of 260, 475, and 725 kbps with UDP congestion

Table 2. Summary statistics for VBR streaming

QP	Mean rate (kbps)	Original avg. PSNR (dB)	Avg. PSNR (dB) at receiver	Avg. end-to-end video frame delay (s)
28	725	36.60	24.33	0.735
36	260	30.25	30.25	0.017
QPver	**475**	**33.44**	**33.20**	**0.199**

[1] Video frame delay was found by finding the time that the latest slice-bearing packet of that frame arrived.

Table 3. Summary statistics for CBR streaming

QP	Mean rate (kbps)	Original avg. PSNR (dB)	Avg. PSNR (dB) at receiver	Avg. end-to-end video frame delay (s)
28	725	36.32	24.22	0.648
36	260	29.83	29.83	0.017
QPver	475	33.36	33.36	0.053

We also considered the situation if the CBR congesting stream was transported through TCP rather than UDP. Essentially, the CBR source is now congestion controlled through the TCP saw-tooth-like [15] restraint mechanism. In Fig. 6, the main impact is upon the higher-quality unconstrained VBR stream, which has sharper peaks and troughs in quality. No doubt this is a result of the TCP source restraining its rate as a result of dropped packets at the IEEE 802.11 AP buffer, resulting in drop bursts when the TCP-transported source turns on again at a 'saw-tooth' peak. Fig. 7 shows the impact upon the CBR encoded video streams. Again the higher–rate video stream now experiences accentuated fluctuations in video quality, due to recurrent TCP transported surges in rate.

4. CONCLUSION

This paper proposes that quality-adaptive VBR, otherwise known as 'feedback VBR' is preferable for wireless network streaming and has demonstrated this in a WLAN home network. The proposed scheme in the home network used local adaptation of the VBR rate through bitrate transcoding, though scalable approaches are also possible. It was found that unconstrained VBR resulted in weaker or unacceptable quality video, when a CBR data source competed for 802.11 AP buffer space while

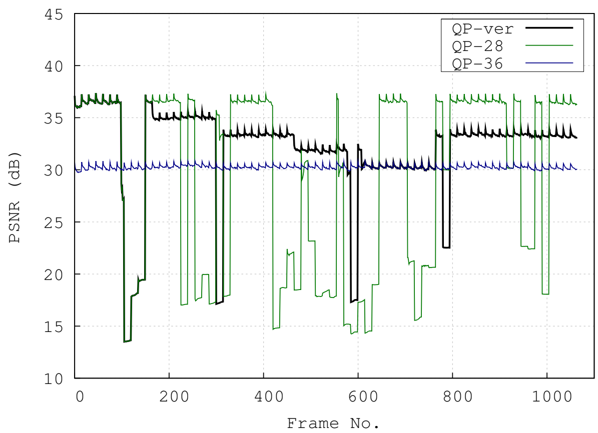

Fig. 6. VBR-encoded objective video quality (Y-PSNR) for HN video streaming at QP 28 and 36, and variable QP (QPver) with TCP congestion

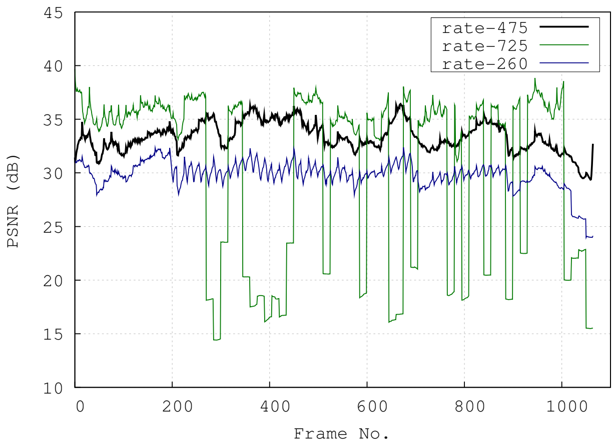

Fig. 7. CBR-encoded objective video quality (Y-PSNR) for HN video streaming at rates of 260, 475, and 725 kbps with TCP congestion

streaming from a home NAS unit. This was particularly the case when UDP transport was used by the congesting source, though TCP file downloads also pose a threat to unconstrained VBR and CBR streaming alike. Future work will directly compare native streaming through VBR quality adaptation, when streaming either from a remote source or when an HTTP adaptive streaming (through progressive download) source has been temporarily stored on a home network. From these various results, a distortion cost model that links VBR, the use of buffering, and PSNR can be elaborated.

REFERENCES

[1] T. Stockhammer, "Dynamic adaptive streaming over HTTP–design principles and standards," in *Proc. ACM Conf. on Multimedia Systems*, 2011, pp. 133-144.

[2] L.D. Cicco, S. Mascolo, P. Bari, and V. Orabona, "An experimental investigation of the Akamai adaptive video streaming," in *Proc. Int'l Conf. on HCI in Work and Learning, Life and Leisure.* 2010, pp. 447-464.

[3] R. Razavi, M. Fleury, and M. Ghanbari, "Unequal protection of video streaming through adaptive modulation with a tri-zone buffer over Bluetooth Enhanced Data Rates," *EURASIP J. on Wireless Commun. and Networking,* [online volume] 16 pages, 2008.

[4] D. Lefol, D. Bull, and N. Canagarajah, "Performance evaluation of transcoding algorithms for H.264," *IEEE Trans. Consum. Electron.*, vol. 52, no. 1, pp. 215-222, 2006.

[5] S. Notebaert, J. De Cock, and R. Van de Walle, "Improved H.264/AVC requantization transcoding using low-complexity interpolation filters for 1/4-pixel motion compensation," in *Proc. IEEE Symp. on Computational Intelligence in Signal Processing*, 2007, pp. 317-312.

[6] T. Schierl, T. Stockhammer and T. Wiegand, "Mobile video transmission using Scalable Video Coding (SVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1204-1217, 2007.

[7] T.V. Lakshman, A. Ortega, and A.R. Reibman, "VBR video: Tradeoffs and potentials," *Proc. of the IEEE*, vol. 86, no. 5, May 1998, pp. 952–973.

[8] X. Kang, J. Lan, L. Liu, and X. Zhuang, "SNR-based bit allocation in video quality smoothing," in *Proc. Pacific Rim Conf. on Multimedia*, 2006, pp. 289-998.

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

[10] W. Feng and J. Rexford, "Performance evaluation of smoothing algorithms for transmitting prerecorded variable-bit-rate video," *IEEE Trans. Multimedia*, vol. 1, no. 3, pp. 302–312, 1999.

[11] J. McManus and K. Ross, "Video on demand over ATM: Constant-rate transmission and transport," in *Proc. IEEE INFOCOM*, 1996, pp. 1357–1362.

[12] G. Bhatti, Z. Shinoglu, K. A. Peker, J. Guo, and F. Matsubara, "A TV-centric home network to provide a unified access to UPnP and PLC domains," in *IEEE Int'l Workshop on Networked Devices*, 2002, pp. 234-242.

[13] T.S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 2009.

[14] Q. Huynh-Thu, and M. Ghanbari, "The accuracy of PSNR in predicting video quality for different scenes and frame rates," *Telecommunication Systems*, vol. 49, no. 1, pp. 35-48, 2012.

[15] J.F. Kurose, and K.W. Ross, *Computer networking: A top-down approach*, 5th ed., Pearson, Boston, MA, 2009.