CONGESTION-RESISTANT SCALABLE MEDIA STREAM MAPPING FOR AN IEEE 802.11E SENSOR NETWORK

Ismail A. Ali, Martin Fleury, and Mohammed Ghanbari

University of Essex, United Kingdom

ABSTRACT

Mapping an H.264/SVC to the IEEE 802.11e video service class may result in the decoder discarding an excessive number of packets. This is because of dependencies between the discarded packets and those dropped in the service queue. Instead, the paper proposes mapping SVC quality layers across multiple service classes, thereby reducing the impact of traffic congestion upon video quality. Video quality in the sensor network scenario tested was almost stabilized despite up to 10% packet drops through traffic congestion. An added advantage of the mapping is that the delay of the base-layer packets was considerably reduced compared to the standard IEEE 802.11e mapping.

Index Terms**—**ad hoc wireless network**,** H264/SVC, IEEE 802.11e EDCA

1. INTRODUCTION

By more effectively exploiting IEEE 802.11e Enhanced Distributed Channel Access (EDCA) [1], we propose a congestion-resistant media stream mapping for the Scalable Video Coding (SVC) extension [2] to the H.264/AVC (Advanced Video Coding) codec. When it is applied to an IEEE 802.11b wireless network operating in ad hoc (infrastructure-less) mode, the mapping reduces the deterioration of received video quality in the face of increased traffic congestion. A further advantage of the mapping is that it is possible to trade per-layer end-to-end latency for video quality. This is because in this mapping scheme successive enhancement layers are delayed relatively more in comparison to the base layer. Thus, higher-quality video is possible for one-way streaming (with a small increase in start-up delay) but lower-quality base-layer video, which is compatible to H.264/AVC, can be opted for in an interactive service such as video phone. The scenario explored in this paper is a sensor network, possibly operating in an emergency, in which differing quality video from a remote surveillance camera is extracted depending on the interest of the scene to the monitor.

IEEE 802.11e Enhanced Distributed Channel Access (EDCA) adds quality-of-service (QoS) support to legacy IEEE 802.11 wireless networks by introducing four Access Categories (ACs), viz. AC0, AC1, AC2, and AC3 for background (BK), best-effort (BE), Video (Vi) and Voice (Vo) respectively in order of increasing priority. However, this paper shows that mapping different SVC layers across the 802.11e EDCA ACs is an effective alternative to placing the complete SVC stream into AC2, because it helps to counter inter-layer dependencies and to preserve reference pictures typically found in the base quality layer, without which other packets may be discarded by the decoder because of these dependencies. By reducing the discard rate the quality of delivered video is almost stabilized within the range of interest despite increasing packet drop percentages. Moreover, the mapping has the advantage that it allows selective recovery based on a delay criterion.

Perhaps, the nearest prior work to ours is that in [3], when priority sets and partitioned coding data from the non-scalable H.264/AVC codec were mapped to different EDCA ACs to improve the quality of single-layer video. Work in [4] also combined SVC with IEEE 802.11e EDCA but proposed packet aggregation to reduce contention overhead arising from many small SVC packets but kept the same AC for the entire stream. A generalized architecture for dynamically mapping H.264/SVC to wireless QoS class appeared in [5] though the results presented did not examine the video performance in detail. While no specific modifications for video streaming were made, an interesting possibility [6] is to dynamically adjust access priorities depending on channel conditions, especially expected latency.

It is important to notice [7] that SVC's SNR rate-distortion performance is only about 0.5 dB below or 10% of the rate of single-layer H.264/AVC. The paper now further explains the background to the proposed mapping scheme.

2. BACKGROUND

A scalable bit-stream is one in which parts of the stream can be removed but the remainder is still decodable. An H.264/SVC bit-stream consists of a base layer and one or more nested quality enhancement layers. In the SVC extension scalability across the Signal-to-Noise Ratio (SNR) (quality) layers is achieved by hierarchical and inter-layer reconstruction. Though this form of reconstruction improves coding efficiency considerably it introduces dependencies between and within layers, which is a potential weakness which this paper seeks to mitigate.

In coarse-grain scalable (CGS) quality layers, residual textural information (transform coefficients) from the prior layer is re-quantized at a finer resolution but there is a loss in rate switching flexibility [7] as switching may only occur at IDR pictures. However, Medium-Grained Scalability (MGS) adopted herein is additionally supported at the Network Abstraction Layer unit (NALU) level, a NALU being a logical H.264 packet with RTP header. Because it does not support predictive reference, progressive fine-grained scalability (FGS) is not supported in SVC. In fact, FGS would significantly decrease coding efficiency compared to single-layer H.264/AVC, for comparable rate-distortion performance to SVC [7]. Though MGS improves coding efficiency it introduces further coding dependencies, which can cause the decoder to discard packets if the dependent packets have been dropped from the EDCA queues.

Fig. 1 shows an illustrative size four SVC Group-of-Picture (GOP) structure with hierarchical B-pictures (or P-pictures) between the key frames. In tests a GOP size of 16 was employed. We have also employed key pictures [2] which form the coarsest temporal layer and which serve to delimit the extent of drift within motion-compensation decoding.

Turning to IEEE 802.11e EDCA, Fig. 2 shows the associated queues (set to 40 variable-sized packets in tests) with entry to the queues defined by a mapping function. Should several packets emerge simultaneously from the queues then contention is resolved by the virtual collision handler before a transmission attempt. Each AC has different Distributed Coordination Function (DCF) parameters for the Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) back-off mechanism. In tests (Section 3) the IEEE 802.11e MAC parameter values [1] for the IEEE 802.11b radio were employed but an extension is to tune these parameters to set a desired quality/delay trade-off. The parameters include Contention Window (CW) minimum and maximum sizes and arbitrary inter-frame space (AIFS). The effect is to delay access to the channel from numerically lower AC queues, thus increasing the probability that packets from these queues will be dropped through buffer overflow due to arriving packets. As IEEE 802.11b was in ad hoc mode, the well-known Ad-hoc On-Demand Distance Vector (AODV) routing protocol was deployed with control packets carried at higher-priority AC3.

3. SVC TO EDCA MAPPING

The proposed mapping between SVC video traffic and EDCA ACs can take many forms depending on the desired types and levels of scalability. In Fig. 3, base layer video is assigned to AC2, as this is the default AC for video traffic. We mapped enhancement layer traffic to AC1 and AC0,

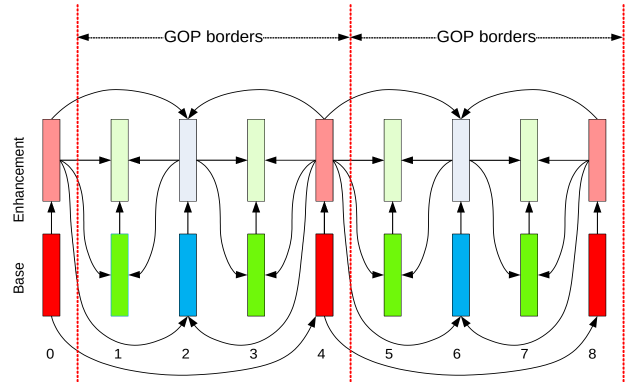

Fig. 1. Illustrative GOP coding structure for GOP size of four with base and single enhancement layer, showing the prediction structure. Pictures appearing at 0, 4, and 8 are coded as key pictures.

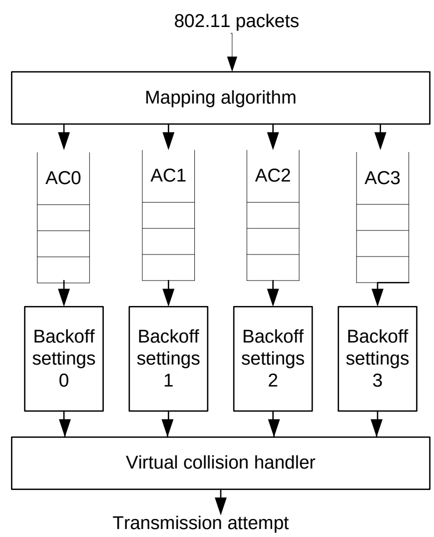

Fig. 2. IEEE 802.11e Medium Access Control (MAC) layer architecture.

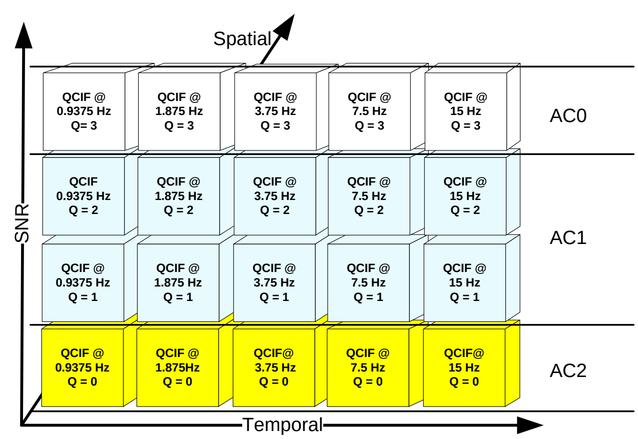

Fig. 3. Possible H.264/SVC stream mapping to IEEE 802.11e ACs, Q is quality level identifier.

reserving AC3 for audio and control packets. The spatial resolution, Quarter Common Intermediate Format (QCIF)—176 × 144 pixel/frame, remained constant. Five extractable temporal layers were embedded at 15 Hz (frame/s) right down to about 1 Hz slow scan. The paper now evaluates the impact of the proposed mapping.

4. EVALUATION

We employed the IEEE 802.11e EDCA MAC model developed by the Technical University of Berlin for the well-known ns-2 simulator. For each test, more than 200 independent runs with differing packet drop percentages were plotted. To analyze the performance, we applied our mapping to the static scenario shown in Fig. 4, with all adjacent nodes 200 m apart. The IEEE 802.11b physical layer was modeled with a transmission range of 250 m and a data-rate of 1 Mbps, as appropriate for a sensor network.

We compared the luminance (Y)-PSNR versus percentage packet drop by changing the Constant Bitrate (CBR) traffic rate, assigned to AC1, from node 4 to node 6 (with UDP transport) between 5 and 70 kbps and varying the percentage of Uniformly distributed errors inserted from 0 to 5%. CBR packet size was 1 kB. An FTP stream, assigned to AC0, was sent from node 5 to node 7 with TCP transport and packet size 500 B. As an input, 500 frames of the 'Paris' sequence were encoded with the JSVM v. 9.19.1 encoder. Paris consists of two figures seated around a table in a TV studio setting, with high spatial-coding complexity. For comparison with other results, previous frame replacement was used as a simple form of error concealment. The video packet size was limited to a maximum of 1 kB. One SNR enhancement layer was used with transform coefficients written to three MGS layers and, hence, a finer scalable granularity was achieved. The Quantization Parameters (QP) was set to 35 and 31 for the base and enhancement layer respectively. The motion estimation and mode decision QP for the base layer was set to 34 and to 30 for the enhancement layer. Consequently, the composite SVC average bit-rate was 69.8 kbps, with up to layer 2, 1, and base layer amounting to 58.3, 47.0, and 31.4 kbps respectively, corresponding (before packet drops) to mean YPSNRs of 36.2, 35.3, 34.6 and 33.6 dB.

Fig. 5 shows video quality when the SVC video stream is entirely assigned to AC2. Notice in this Figure and others that, because the packets of two quality layers are mapped to the AC1 queue, video quality is assessed up to layer 2, rather than separate assessments for the two quality layers. As can be seen there is a steady decline in quality as the packet drop percentage increases. Of course, the packet drop percentage is a reflection of network congestion. However, if the mapping of Fig. 3 is applied, then the video quality declines at a much slower rate, Fig. 6. Effectively, the video quality is stabilized over the packet drop range of interest. The reason

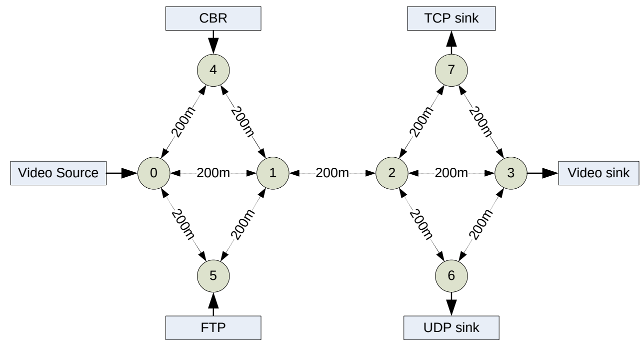

Fig. 4. Simulated IEEE 802.11b test scenario.

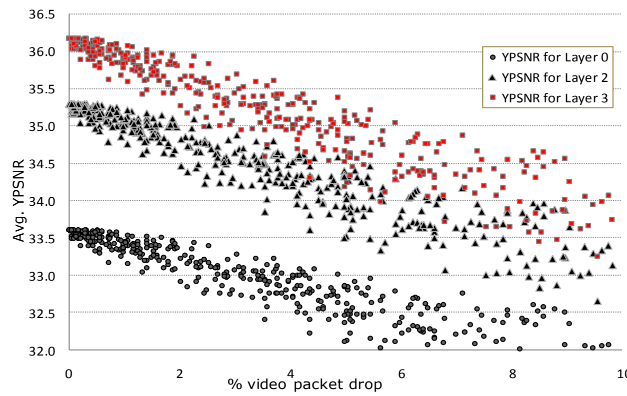

Fig. 5. Video quality (YPSNR) after allocation of SVC stream to AC2 according to reconstruction level.

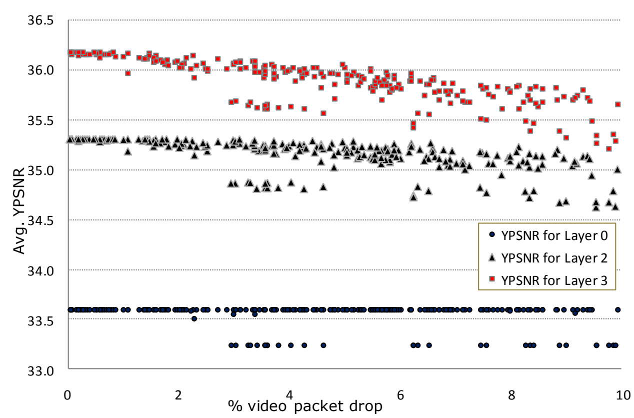

Fig. 6. Video quality (YPSNR) after allocation of SVC stream layers to AC0, 1, 2 according to reconstruction level.

for this phenomenon arises not because of packet drops but because of the number of packets that are discarded by the decoder, as packets upon which the decoder relies to reconstruct others have been dropped. In Fig. 7, the discard rate reaches 25%, though the drop rate is below 10%. However, with protection principally given to the base layer, in Fig. 8, there is reduced self-congestion arising from packets from other quality layers. This results in more base-layer packets surviving and consequently less packets are discarded as not decodable. Turning to end-to-end delay, it

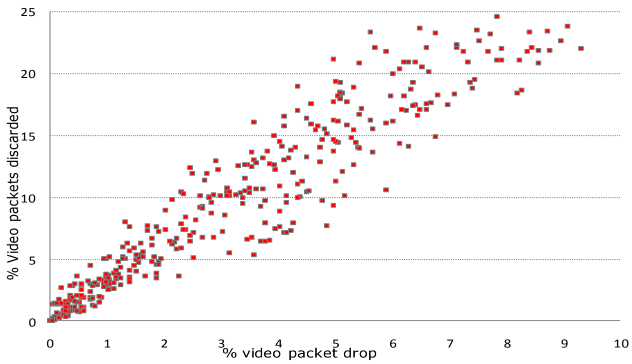

Fig. 7. Composite SVC stream packet discard percentages for allocation to AC2.

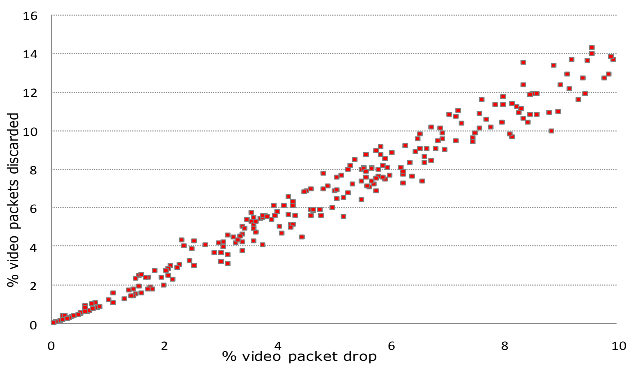

Fig. 8. Composite SVC stream packet discard percentages for allocation across AC0, AC1, and AC2.

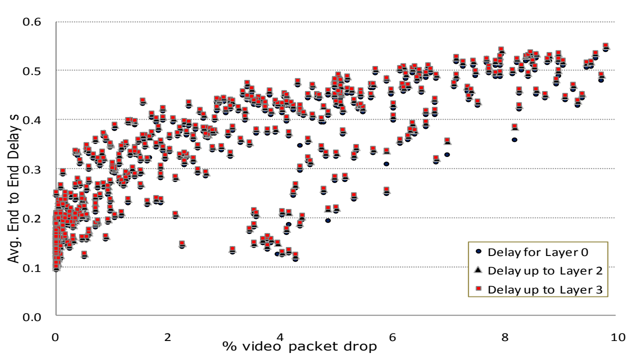

Fig. 9. Average end-to-end video packet delay after allocation of SVC stream to AC2.

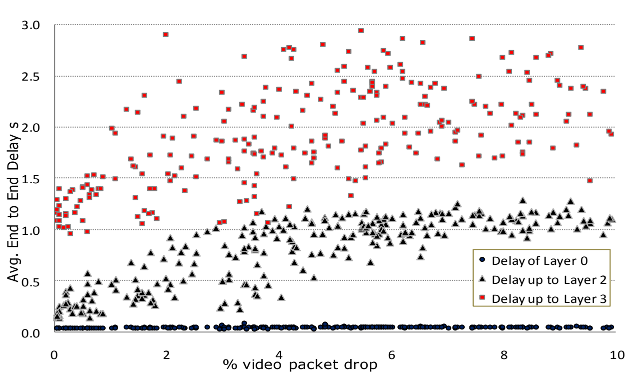

Fig. 10. Average end-to-end video packet delay after allocation of SVC stream layers to AC0, 1, and 2.

is apparent from Fig. 9 that packets from all layers encounter similar levels of delay, which is unsurprising as the packets all share the same queue. However, if the mapping of Fig. 3 is applied, the delay is indeed banded according to layer as in Fig. 10. To reconstruct to full quality, delays of up to 3 s are possible, which is unacceptable for interactive video applications. However, base-layer packets experience less delay than in the EDCA standard mapping, making video at this lower quality (about 33 dB from Fig. 6) possible in some scenarios.

4. CONCLUSION

In our proposal, mapping SVC MGS enhancement layers across several IEEE 802.11e ACs results in up to 2 dB improvement in delivered video quality. This is because in the given scenario the mapping helps to protect packets in the base layer from buffer overflow. Otherwise, even though other packets cross the network without being dropped they cannot be reconstructed by the decoder, as it lacks the dropped packets. Further work will evaluate packetization issues. It will also consider other scenarios in which the latency-weighting that is a product of the mapping can be exploited.

REFERENCES

[1] IEEE 802.11e, “Wireless LAN Medium Access Control and Physical Layer specifications Amendment 8: Medium Access Quality of Service Enhancements,” 2005

[2] H. Schwarz, D. Marpe, and T. Wiegand, “Overview of the scalable video coding extension of the H.264/AVC standard,” *IEEE Trans. Circ. Syst. Video Technol*., vol. 17, no. 9, pp. 1103–1119, 2007.

[3] A. Ksentini, M. Naimi, and A. Guéroui, “Toward an improvement of H.264 video transmission over IEEE 802.11e through a cross-layer architecture,” *IEEE Communs. Mag.*, vol. 44, no. 1, pp. 107–114, 2006.

[4] A. Fiandrotti, D. Gallucci, E. Masala, and J. C. De Martin, “High-performance H.264/SVC video communications in 802.11e ad hoc networks,” *First Euro-NF Workshop*, pp. 200 – 210, 2008.

[5] J. Huuskoa, J. Vehkapera, P. Amon, C. Lamy-Bergot, G. Panzad, J. Peltolaa, and M.G. Martinie, “Cross-layer architecture for scalable video transmission in wireless network,” *Signal Processing: Image Communication*, vol. 22, pp. 317–330, 2007.

[6] Y.-J. Wu, J.-H. Chiu, and T.-L. Sheu, “A modified EDCA with dynamic contention control for real-time traffic in multi-hop ad hoc networks,” *J. of Inform. Sci. and Eng*., vol. 24, pp. 1065–1079, 2008.

[7] T. Schierl, Y. S. De la Fuente, C. Hellge, and T. Wiegand, “Priority-based transmission scheduling for delivery of scalable video coding over mobile channels,” *ICST Mobimedia workshop EUMOB*, 2009

[8] M. Wein, H. Schwarz, and T. Oelbaum, “Performance analysis of SVC,” *IEEE Trans. Circ. Syst. Video Technol*., vol. 17, no. 9, pp. 1194–1203, 2007.