Bandwidth Efficient Data-partitioned Video Streaming for Broadband Wireless

Laith Al-Jobouri, Martin Fleury, and Mohammed Ghanbari
University of Essex, Colchester, United Kingdom
{lamoha, fleum, ghan}@essex.ac.uk

Abstract—**The proposed scheme in this paper combines adaptive channel coding with additional duplicate packets. Adaptive channel coding takes place by: varying the amount of redundancy according to channel conditions; and by retransmission of extra redundant data to reconstruct corrupted packets. In video streaming, greater bandwidth efficiency occurs by transmitting the same video quality at a reduced bitrate. This paper shows that the proposed protection scheme for data-partitioned video achieves that objective when streaming over an IEEE 802.16e (mobile WiMAX) channel.**

Keywords- data-partitioning; duplicate slices; IEEE 802.16e; Raptor channel coding; video streaming

I. INTRODUCTION

Bandwidth efficiency[1] is a measure of the information rate that a wireless technology can support. Thus, if it is possible to transmit a video stream over the same wireless technology but with greater efficiency then the streaming scheme is commercially attractive. This paper shows that a data-partitioned source coding scheme [1] for equivalent video quality achieves greater bandwidth efficiency when streaming over an IEEE 802.16e (mobile WiMAX) [2] wireless channel than a scheme without data-partitioning. Both schemes protect the video stream with a combination of adaptive Raptor channel coding [3] and slice duplication. (The advantages of Raptor coding over similar channel coding schemes were demonstrated in [4].) CBR video is a normal choice of broadcasters in order to reserve bandwidth and storage for pre-encoded video. However, in the case of the scheme without data-partitioning, the inability to duplicate just part of the stream leads to a decrease in efficiency compared to data-partitioned constant bit-rate (CBR) streaming.

In an H.264/AVC (Advanced Video Coding) codec, when data-partitioning is enabled [5], every slice[2] is divided into three separate partitions in order of decoding priority. These data are packed into different Network Abstraction Layer Units (NALU's). Each NALU encapsulated into an IP/RTP/UDP packet for IP Multimedia Subsystem (IMS) transport. For simplicity of comparison, this work assumes a single slice per frame representation but it is perfectly possible to apply data-partitioning to a multi-slice video frame. When motion-copy error concealment is enabled at a decoder, then receipt of a partition-A carrying packet is sufficient to enable a partial reconstruction of the frame. Error concealment by previous frame replacement, though simpler to implement, can lead to inappropriate replacements if there is active motion between successive frames. Packet loss can arise either through buffer overflow arising from congestion or from a failure of the channel coding decode algorithm (as the Raptor decode algorithm is probabilistic). Thus, by only duplicating partition-A packets it is possible to give a measure of protection against outright packet loss of any of the data-partitioned packets. Without data-partitioning, to achieve the same level of protection, it is necessary to duplicate the complete slice. In simulations, the same channel coding scheme is applied to data-partitioned and non-data-partitioned packets to aid reconstruction of packets that are not lost but corrupted. However, though objective video quality is generally similar the bandwidth efficiency is not, because of the burden of transmitting a complete duplicate slice.

Data-partitioning has less overhead than other forms of resilience such as Flexible Macroblock Ordering [5] and, hence, data-partitioning can operate during favorable channel conditions, as well as unfavorable channel conditions. Our scheme, which was introduced in [6], allows graceful degradation in the face of channel error, and in adverse channel conditions, duplicate partition-A packets are transmitted, as illustrated in the paper. On the other hand, the duplicate partition-A stream should be turned off during favorable channel conditions. In an H.264/AVC codec, it is instead possible to send redundant pictures slices [7], which employ a coarser quantization than the main stream, but this can lead to encoder-decoder drift. Besides, for data-partitioning, replacing one partition with a redundant slice with a different quantization parameter (QP) to the other partitions would not permit reconstruction in an H.264/AVC codec.

II. VIDEO PROTECTION SCHEME

This Section now outlines the proposed video streaming protection scheme.

A. Data-partitioning

When H.264/AVC data partitioning is enabled, every slice is divided into three partitions and each partition is located in either of type 2 to type-4 NALU's, as listed in Table I. A NALU of type 2, also known as partition-A, comprises the most important information of the compressed video bit stream of P- and B-pictures, that is the macroblock (MB) addresses, motion vectors (MV's) and essential headers. If any MB's in

---

[1] We have borrowed this term from the physical layer measure of bits/Hz otherwise known as spectral efficiency.

[2] A slice is a sub-frame coding unit provided with decoder resynchronization points.

these pictures are intra-coded, their transform coefficients are packed into the type-3 NALU, also known as partition-B. Type 4 NAL, also known as partition-C, carries the transform coefficients of the motion-compensated inter-picture coded MB's. Partition A is unaffected by content type and, hence, it occupies a smaller percentage of the datarate for higher quality video. Partition-B is generally small but its size is content dependent. Partition-C size can be large for broadcast quality video but for streaming to a mobile device its size may be similar to partition-A's size,

Fig. 1 shows that, of four common error resilience tools in H.264/AVC [5], data partitioning has the least overhead. The illustration is for the well-known *Foreman* clip, representing the jerky motion of a hand-held camera with a rapid pan towards the end of the sequence. In Fig. 1, the horizontal axis represents the mean bitstream rate arrived at by setting the QP to the given value, while the vertical axis represents the mean overhead rate with that QP. As the quality decreases (higher QP), the advantage of data-partitioning increases, as the relative overhead of all schemes increases. In order to decode partition-B and -C, the decoder must know the location from which each MB was predicted, which implies that partitions B and C cannot be reconstructed if partition-A is lost.

Though partition-A is independent of partitions B and C, Constrained Intra Prediction (CIP) should be set in the codec configuration [8] to make partition-B independent of partition-C. By setting this option, partition-B MBs are no longer predicted from neighboring inter-coded MBs. This is because the prediction residuals from neighboring inter-coded MBs reside in partition-C and cannot be accessed by the decoder if a partition-C packet is lost. There is a by-product of increasing packet sizes due to a reduction in compression efficiency but the increase in size may be justified in error-prone environments.

B. *Channel coding*

Prior use of application-layer Raptor coding has been block-based rather than byte-based, resulting in longer repair latencies. The wireless standards also do not consider channel coding rate adaptation. To achieve adaptation, channel estimation is necessary. In fact, the IEEE 802.16e standard specifies that a mobile station should provide channel measurements, which can either be Received Signal Strength Indicators or may be Carrier-to-Noise-and-Interference Ratio measurements made over modulated carrier preambles.

In the proposed adaptive scheme, the probability of channel byte loss through small-scale fading (*BL*) serves to predict the amount of redundant data to be added to the payload. The instantaneous *BL* serves to calculate the amount of redundant data adaptively added to the payload. In an implementation, *BL*, is found through measurement of channel conditions. If the original packet length is *L*, then the redundant data is given simply by:

$$R = L\times BL + (L\times BL^2) + (L\times BL^3)\ldots \\ = L/(1-BL), \quad (1)$$

which adds successively smaller additions of redundant data, based on taking the previous amount of redundant data multiplied by *BL*.

TABLE I. NALU TYPES

NAL unit type	Class	Content of NAL unit
0	-	Unspecified
1	VCL	Coded slice
2	VCL	Coded slice partition-A
3	VCL	Coded slice partition-B
4	VCL	Coded slice partition-C
5	VCL	Coded slice of an IDR picture
6-12	Non-VCL	Suppl. info., parameter sets, etc.
13-23	-	Reserved
24-31	-	Unspecified

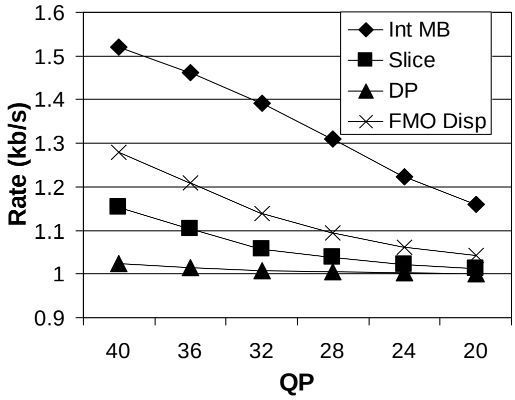

Fig. 1. QCIF *Foreman* rates according to QP (horizontal axis) plotted against overhead rate (vertical axis) arising from H.264 error resilience tool: Int MB = Intra-coded macroblock refresh (one line), FMO Disp = Flexible Macroblock Ordering dispersed mode (two slices), DP = data-partitioning, Slice = slice structuring with 3 slices per frame.

Rateless code decoding in traditional form operates by a belief-propagation algorithm [3] which is reliant upon the identification of clean symbols. This latter function is performed by PHYsical-layer FEC, which passes up correctly received blocks of data (through a cyclic redundancy check) but suppresses erroneous data. For example, in IEEE 802.16e, a binary, non-recursive convolutional encoder with a constraint length of 7 and a native rate of 1/2 operates at the PHY layer.

III. EVALUATION METHODOLOGY

In this Section the adaptive channel coding scheme is described, along with channel modeling and details of the IEEE 802.16e simulation configuration.

A. *Adaptive Raptor coding model*

The following statistical model [9] allowed emulation of Raptor coding 's essential properties:

$$P_f(m,k) = 1 \text{ if } m < k, \\ = 0.85\times 0.567^{m-k} \text{ if } m \ge k, \quad (2)$$

where P_f (*m*, *k*) is the decode failure probability of the code with *k* source symbols if *m* symbols have been successfully received (and 1 - P_f is naturally the success probability). The authors of [9] remark and show that for $k > 200$ the model almost perfectly models the performance of the code, which implies that, if blocks are used, approximately 200 blocks should be received before reasonable correction behavior takes place. This observation motivated the choice of bytes within a packet as the symbols, to reduce latencies. Upon receipt of the correctly received data, decoding of the information symbols is attempted, which will fail with a probability given by (2) for k > 200.

If a packet cannot be decoded, despite the provision of redundant data then extra redundant data are added to the next packet. In Fig. 2, packet X is corrupted to such an extent that it cannot be immediately decoded. Therefore, in packet X+1 some extra redundant data are included up to the level that decode failure is no longer certain.

It is implied from (2) that if less than k symbols (bytes) in the payload are successfully received then a further $k - m + e$ redundant bytes can be sent to reduce the risk of failure. This reduced risk arises because of the exponential decay of the risk that is evident from equation (2) and which gives rise to Raptor code's low error probability floor.

In simulations, the decision on whether the Raptor code belief-propagation algorithm would run to completion was by comparing a Uniformly-distributed random variable's value with that of the probability given by (2). In the tests, e was set to four, resulting in a risk of failure (from (2)) of 8.7 % in reconstructing the original packet, if the extra redundant data successfully arrives. If after sending extra redundant data the packet could still not be reconstructed it was declared lost. Similarly if for m > k it was not possible to reconstruct a packet, it too was declared lost.

B. Wireless channel model

In the proposed adaptive scheme, the probability of channel byte loss (*BL*) through in the well-known Gilbert-Elliott channel model serves to predict the amount of redundant data to be added to the payload. There are two hidden states (good and bad) in this Markov model. If P_{GB} and P_{BG} are the probabilities of going from good to bad state and from going from bad to good state respectively, then

$$\pi_G = P_{BG}/(P_{BG} + P_{GB}) \tag{3}$$

$$\pi_B = P_{GB}/(P_{BG} + P_{GB}) \tag{4}$$

are the steady state probabilities of being in the good and bad states. P_G and P_B are the probabilities of (byte) loss in the good and bad states respectively. Both states were modelled by a Uniform distribution. Consequently, the mean probability of channel byte loss is given by

$$BL_{mean}= P_G.\ \pi_G + P_B.\ \pi_B \tag{5}$$

which was the mean of a Uniform distribution. Thus, in simulations *BL* was selected from a Uniform distribution with mean BL_{mean}.

Assuming perfect channel knowledge of the channel conditions when the original packet was transmitted establishes an upper bound beyond which the performance of the adaptive scheme cannot improve. However, we have included measurement noise into the estimate of *BL* to test the robustness of the scheme. Measurement noise was modelled as a zero-mean Gaussian (normal) distribution and added up to

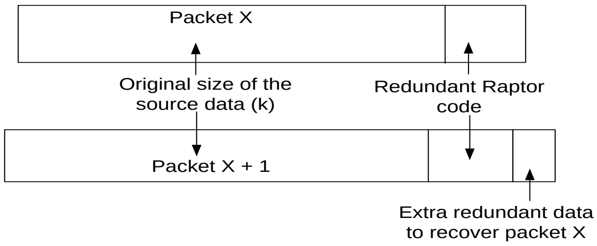

Fig. 2. Division of payload data in a packet between source data, original redundant data and piggybacked data for a previous erroneous packet.

a given percentage (5% in the evaluation) to the packet loss probability estimate.

In the Gilbert-Elliott model for the ns-2 simulator, the settings for fast fading were P_{GG} was set to 0.95, P_{BB} = 0.96, P_G = 0.02 and P_B = 0.165. Each data point is the mean of ten independent runs after the simulator was allowed to reach a steady state.

C. Simulation configuration

The tests were performed *Football* (high temporal coding complexity) video sequence encoded in Common Intermediate Format (CIF) @ 30 Hz. The GoP structure was IPPP…, i.e. one initial I-picture followed by all P-frames. 2% intra-refresh (IR) MBs were randomly inserted to arrest temporal error propagation upon loss of a P-frame. CIP was also set to make partition-B independent of partition-C.

The video stream was transmitted to an IEEE 802.16e mobile station (MS) and, to introduce sources of traffic congestion, a permanently available FTP source was introduced with TCP transport to a second MS. Likewise, a CBR source with packet size of 1000 B (the WiMAX maximum transfer unit) and inter-packet gap of 30 ms was also downloaded to a third MS. Buffer sizes were set at 50 packets to avoid unnecessary delay and reduce energy consumption at the mobile devices. The simulations adopted the mandatory settings for a 10.67 Mbps downlink (DL) rate with 3:1 DL/UL sub-frame ratio for the only WiMAX forum frame size of 5 ms, 16-QAM ½ modulation over a 10 MHz channel with IEEE 802.16e recommended antenna heights and transmit/receive powers.

IV. EVALUATION

The potential savings in bandwidth efficiency for *Football* are guided by Table II. For example, at 1 Mbps, duplication of partition-A adds an extra 37% to the datarate, while duplicating a complete slice obviously adds an extra 100% to the datarate. Space does not permit reporting results for a range of video content or genre. Because Football is more complex to code its partition-A contribution is smaller but conversely the impact of packet loss or corruption is higher because of predictive dependency in the video coding process.

Fig. 3 shows the number of packets dropped outright when streaming *Football* with and without data-partitioning. In the case of duplication, a packet is lost when no duplicate is available. For data-partitioning streaming, obviously more packets are sent. Duplication in all cases reduces the packet drop rate and in the case of streaming at 1 Mbps without data-

TABLE II. PERCENTAGE OF DATARATE FROM EACH PARTITION

Partition:	*A*	*B*	*C*
Football at 500 kbps	52%	14%	34%
Football at 1Mbps	37%	24%	39%

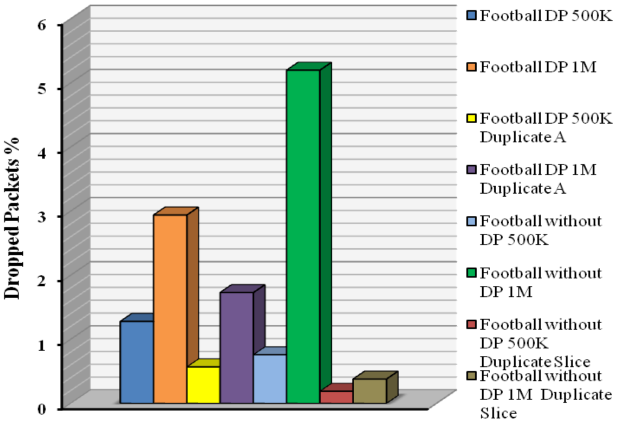

Fig. 3. Dropped packets according to CBR rate, presence of data-partitioning (DP) and/or duplicate slices. 500K = 500 kbps, 1M = 1 Mbps.

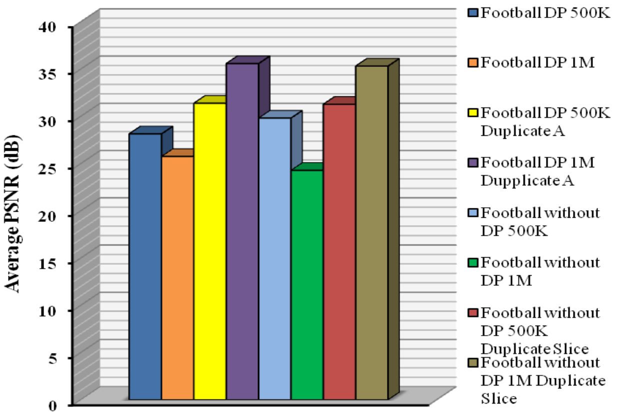

Fig. 4. Y-PSNR according to CBR rate, presence of data-partitioning (DP) and/or duplicate slices. 500K = 500 kbps, 1M = 1 Mbps.

partitioning, the reduction is comparatively large. The gain for the scheme without data-partitioning is relatively larger,though for generally fewer and larger packets. From Fig. 4, the luminance PSNR is better with duplication than without and greater gain results from using duplication at the higher datarate. Video quality at 25–31 dB is approximately equivalent to an ITU-T recommendation P.910 'fair' rating, whereas above 31 dB it is 'good'. Thus, duplication is needed to pass these quality thresholds. However, the main point is that data-partitioning results in equivalent objective video quality to duplication at a higher data rate.

However, this gain is only achieved with adaptive Raptor channel coding protection, as is apparent for the high levels of packet corruption in Fig. 5. Corrupt packets are packets that are potentially repairable by channel coding but may require extra redundant data. In all but one case, the percentage of packets corrupted is larger for the non-data partitioning scheme though the number of packets sent is larger when

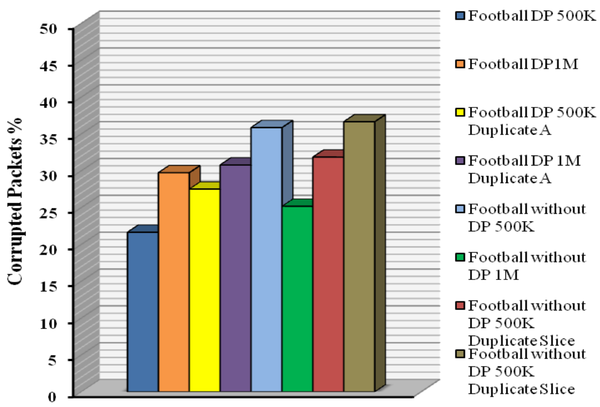

Fig. 5. Corrupted packets according to CBR rate, presence of data-partitioning (DP) and/or duplicate slices. 500K = 500 kbps, 1M = 1 Mbps.

employing data-partitioning. The main impact of packet corruption is delay from retransmission. The non-data-partitioned scheme incurs greater delay (a mean of 4 to 5 ms for 500 kbps and 1 Mbps with duplication compared to 2 ms for the equivalent data-partitioned packets), as well as a greater risk of corruption. Though the delays are generally small, there is the possibility of accumulated delay leading to missed display deadlines for long video streams.

V. CONCLUSION

This paper, employs duplicate partition-A's but equal protection through channel coding. The paper demonstrated that the main gain from this arrangement is more effective bandwidth utilization, though there is a second-order advantage arising from a reduction in latency. Future work should document the gain across different video genres.

REFERENCES

[1] T. Stockhammer, and M. Bystrom, "H.264/AVC data partitioning for mobile video communication," in *IEEE Int'l Conf. on Image Processing*, 2004, pp. 545-548.

[2] IEEE, 802.16e-2005. "IEEE standard for local and metropolitan Area networks. Part 16: Air interface for fixed and mobile broadband wireless access systems," 2005.

[3] A. Shokorallahi, "Raptor codes," *IEEE Trans. Info. Theory*, vol. 52, no. 6, pp. 2551-2567, 2006.

[4] P. Palanisamy, and T.V.S. Sreedhar, ""Performance analysis of Raptor codes in WiMAX channel system over fading channel," in *IEEE TENCON*, 2008, 5 pages.

[5] S. Wenger, "H.264/AVC over IP," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 645-655, 2003.

[6] L. Al-Jobouri, M. Fleury and M. Ghanbari, "Raptor coding of H.264 data-partitioned cideo over a WiMAX channel", in *IEEE Int'l Conf. on Consumer Electronics*, 2011, pp. 341-342.

[7] J. Radulovic, Y.-K. Wang, S. Wenger, A. Hallapuro, M.H. Hannuksela, and P. Frossard, "Multiple description H.264 video coding with redundant pictures," in *Int'l Workshop on Mobile Video*, 2007, pp. 37-42.

[8] Y. Dhondt, S. Mys, K. Vermeirsch, and R. Van de Walle, "Constrained inter prediction: Removing dependencies between different data partitions," in *Advanced Concepts for Intelligent Visual Systems*, 2007, pp. 720-731.

[9] M. Luby, T. Gasiba, T. Stockhammer, and M. Watson, "Reliable multimedia download delivery in cellular broadcast networks," *IEEE Trans. Broadcast.*, vol. 53, no. 1, pp. 235-246, 2007