Enhanced Slicing for Robust Video Transmission

Sandro Moiron, Ismail Ali, Mohammed Ghanbari, Martin Fleury
School of Computer Science and Electronic Engineering
University of Essex, Colchester, United Kingdom
{smoiro, iaali, ghan, fleum}@essex.ac.uk

***Abstract*—This paper presents a low complexity yet effective slicing scheme for enhanced error resiliency. The proposed scheme uses Flexible Macroblock Ordering (FMO) without requiring any decoder modification without any increase in bitrate or computational complexity. Frame level packet priority is then provided based on slice position within the frame. Experimental results indicate that, when FMO is combined with Unequal Error Protection, the scheme can achieve quality gains up to 4 dB over other slicing techniques.**

I. Introduction

In the last decade the increasing user demand for new mobile multimedia services and applications has lead to the development of better video compression standards and wireless technologies. The H.264/AVC codec [1] was developed with low bandwidth streaming applications in mind, aiming to achieve a high compression efficiency.

As the video compression efficiency improves, each bit carries more information becoming more error sensitive which results in a large quality degradation in error prone channels. When transmission errors occur, pixels represented in those packets become non decodable and are replaced by concealed versions of the previously received pixels. In order to combat against transmission errors, the H.264/AVC video compression standard has considered several error resilient techniques such as Flexible Macroblock Ordering (FMO), Data Partitioning (DP), Redundant Slices (RS) [2].

In order to achieve high compression ratios, most video codecs exploit temporal redundancy between the frames. The H.264/AVC can use up to 16 previously coded frames as references for the current frame being encoded (known as inter frame prediction). This not only increases the compression efficiency but also the inter frame dependency. Hence, errors in a distorted frame will propagate to the following frames as error concealment may not clear whole distorted parts.

To prevent temporal error propagation intra coded macroblocks are normally used to refresh the damaged area. Intra coding exploits only spatial redundancy within the frame removing therefore any inter frame dependency. Intra refresh can be applied in two different ways. In one method the complete frame is intra coded (I-frame). This technique is the most effective at mitigating error propagation as it breaks all dependencies from the previously coded frames. However, the reduced compression efficiency of intra coded frames, compared with temporally predicted frames, makes them not suitable for continuous use. Instead it is better to use mainly the highly compression efficient inter coded frames with some periodic intra coded frames to limit potential error propagation that might occur. Even so, when sending video over low bandwidth channels, intra frames introduce bursts of high bitrates often undesirable for low channel rates. In order to prevent error propagation and at the same time remove bursts of high bitrates a group of blocks in each frame is intra refreshed. This is normally done in a cyclic manner such that after a number of frames the whole frame is updated.

In cyclic intra refreshing a line of MB in each frame, some parts of the frame are more likely to encounter higher error propagation than others. For example, those macroblocks that are refreshed in a frame are less likely to be erroneous in the following frame. At the same time, since they were just refreshed, if they face errors in the next few frames they are more likely to stay erroneous for a long period, increasing the potential number of erroneous reference frames. On the other side, macroblocks to be refreshed in the next frames represent a small threat to error propagation as they are to be intra coded shortly. Hence, it is clear that there is an unequal importance of those macroblocks above and below the intra refresh line.

In [3][4][5], the strategy was to use a adaptive intra refreshing scheme which required the encoder to keep track on which parts of the image area were recently refreshed. The encoder would then refresh those macroblocks which were more subject to error propagation. Other schemes [6][7] were also proposed where 'explicit' FMO has been employed with adaptive macroblock grouping. The grouping schemes were based on individual MB importance measuring the quality impact when an individual MB is lost. These methods increase significantly the computational complexity due to implicit video content analysis for the macroblock rating. Additionally, methods using 'explicit' FMO also increase the bitrate to transmit the updated macroblock maps for every frame.

This paper presents a new slicing scheme for enhanced error resiliency of video encoded with cyclic intra refresh lines. Unlike other techniques, the proposed scheme allows unequal error protection without introducing any additional computational complexity or increase in bitrate. Experimental results show that a significant quality gain up to 4 dB can be achieved compared with other slicing techniques.

The paper follows in Section II with an introduction to error resilience and slicing techniques. Section III describes the proposed slicing method illustrating the advantages and potential application scenarios for its usage. Experimental results are presented and a comprehensive analysis is carried out in Section IV. Finally Section V completes the paper with some concluding remarks.

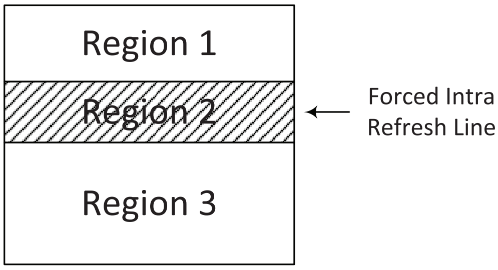

Fig. 1. Proposed Slicing Scheme

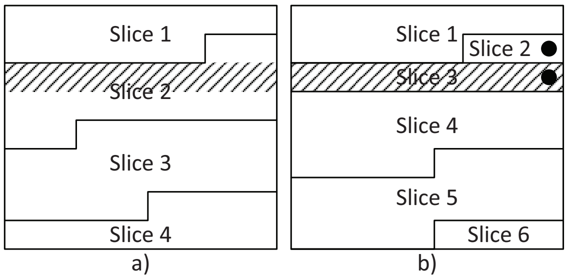

Fig. 2. Classic Slicing (left) vs Enhanced Slicing (right)

II. Slicing and its Implications in Error Robustness

The ever growing need to deploy video content to numerous new applications and networks has raised the question on how to handle this wide variety of combinations [8]. As a result, the H.264/AVC was designed to separate video compression from network delivery. Hence, all video related tasks are handled by the Video Coding Layer (VCL) while network adaptation issues are handled by the Network Abstraction Layer (NAL). VCL defines all the coding structure to efficiently represent the video content and forward it to the NAL. The NAL then is responsible for the adaptation to the underlying transport layer. The content coming from the VCL layer is encapsulated into NAL units (NALu). A NALu consists of a one byte header followed by the corresponding payload information. In the H.264/AVC codec, each frame is decomposed into macroblocks which are then grouped into slices. Every slice is independently decodable being therefore considered as re-synchronization points that prevent error propagation to the entire picture.

Depending on the type of transmission network and the network technology in use the NAL packet size is limited to the maximum transmission unit (MTU) supported by the network. If an error cannot be corrected in a packet, the whole packet is discarded, resulting into a significant loss of data. Limiting the packet size is therefore regarded as a technique for limiting the amount of data loss. Packetization can be performed either by the network or by the encoder. When performed by the network the bitstream is split into packets to conform with the supported MTU. This process breaks video slices into multiple packets, creating inter packet dependency. As a result, network packetization significantly increases the error sensitivity of the bitstream and is not recommended. Instead, packetization is normally implemented at the encoding stage such that each packet contains a single slice with a size less than the MTU of the underlying network.

As mentioned in the introduction cyclic intra refresh is preferred over the whole frame update. However a side effect in cyclic intra refresh lines is an implicit division of the frame into three regions in terms of potential for error propagation as seen in Figure 1. The first region is located above the intra refresh line and it is the most error sensitive region. This is due to the fact that this regions is recently refreshed with intra coded blocks and will be used as a reference for the upcoming frames without being refreshed with intra coded blocks soon. As a result, an error affecting this region propagates for a long period introducing a significant quality loss. The second region is the intra refresh line itself. In our previous work [9] it was found that although the intra refresh line plays an important role at mitigating error propagation, it uses a significant portion of the bitrate share while still representing a small portion of the image area (22 MBs of the 396 total MBs in a CIF frame). Therefore, it was concluded in [9] that if packets are to be discarded, then intra refresh packets are the ones which introduce the least quality loss for the same percentage of loss rate. Finally, the last region of the frame is located below the intra refresh line. This region has the smallest potential error impact because an error in this region does not propagate for long. If the intra refresh line is cycling horizontally, from top to bottom, those macroblocks below the intra refresh line are to be intra coded in the following frames. Hence, even if errors occur in this region they are to be cleaned next, limiting the number of frames to be affected by the prediction from the distorted reference frames. Overall, it can be said that cyclically intra refreshed video has three regions with unequal impact on error propagation.

To exploit this unequal error sensitivity of the cyclically intra refreshed frames, it is logical that the three parts of a frame to be treated differently. In slicing the frames considering the maximum packet size criteria, it is likely that some slices contain macroblocks from more than one image region as illustrated in figure 2a. The challenge here is how to slice the image such that all three image regions can be separated and differentiated during transmission in order to offer different protection levels. The next section will describe the proposed slicing scheme along with the technique that will allow differentiating each packet. It should be empathized that this slicing scheme does not require any decoder modification as it is compliant with the standard.

III. Proposed Algorithm

In order to slice an image into three distinct regions according to figure 1 the slice structuring of the encoder was modified to accommodate a new set of rules for slicing. These rules prevent the mixture of macroblocks from different regions to be packed into the same slice. They also define when the current slice should be terminated or if more macroblocks should be packed in.

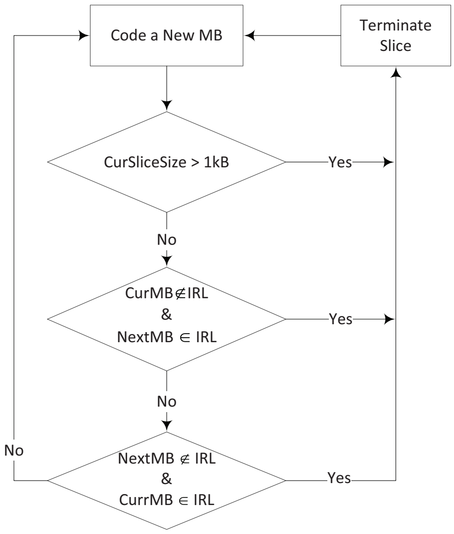

Fig. 3. Robustness aware slicing scheme

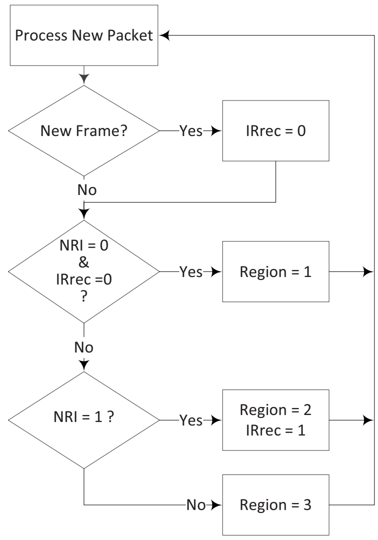

Fig. 4. Identification of different image regions

A. New Slicing Rules

In addition to the original maximum packet size criteria, the first newly introduced rule checks if the current macroblock (CurMB) is the last macroblock before the start of an intra refresh line (IRL) and terminates the slice if this is true. The second new rule checks if the macroblock being encoded is the last macroblock of the intra refresh line, closing the slice like the previous rule if it is true. These positions can be calculated if the frame width and the intra refresh position for the current frame are known in advance. Figure 2b shows the check points (black dots) tested by the two new rules added to the slicing mechanism. Figure 3 shows the diagram for the proposed slicing algorithm to be implemented at the encoder.

B. Network Signaling

During encoding, the encoder is aware of the position of the intra refresh line and is able to implement the proposed slicing scheme. However, after packetizing the video into Real Time Protocol (RTP) packets, intra refresh line packets are no longer identifiable in the bitstream, making it impossible to separate packets according to their associated regions. In order to signal the intra refresh line packets (Region 2) to the network, the priority bits (NRI) in the NAL header are modified. In packetizing slices containing macroblocks from the intra refresh line, the NRI field is set to 1 for easy identification by the network. In all other cases this field is left with the default values.

C. Region Identification on the Network Side

On the network side, it is possible to make a correspondence between the packet and image region by analyzing the RTP TimeStamp, packet number and NRI field. During this process IRrec is a flag which keeps track of whether the IR line was already received in the current frame or not.

In figure 4, the first block checks if the current packet belongs to a new frame by checking the RTP TimeStamp field. If TimeStamps are different (New Frame), then IRrec is reset to signal that no IR was received. Afterwards it performs the checks to identify which region the packet is from. If NRI and IRrec are zero, it means that the current packet does not belong to the IR line and no intra refresh packets were received since the beginning of this frame, thus it belongs to region 1. Else if NRI equals to 1, then the packet belongs to an IR line (region 2) and IRrec is set to 1 to distinguish the following packets between region 1 and region 3. If all conditions fail then the packet is from region 3.

IV. Experimental Results

In order to evaluate the performance of the algorithm hereby presented, the H.264/AVC reference encoder JM16.1[10] was modified to apply the new slicing scheme and the required signaling. The performance was measured in terms of objective quality (PSNR) while simulating a network data drop percentage ranging from 0 to 10%. Six different CIF (352×288 pixels) video sequences (*Akiyo, Bus, Foreman, Mobile, Paris and Stefan*) with 300 frames were encoded with the H.264/AVC 'Main Profile', a single reference frame and CABAC entropic encoding at a constant bitrate of 1 Mbps. RTP packetization mode was used with a maximum packet size of 1kB set at the encoder.

For each sequence, four tests were conducted. In the first test, only packets belonging to the intra refresh line slice were dropped (Drop IR). For the second test, only packets belonging to the region above the intra refresh line were dropped (Drop Above IR) and for the third test those of the region below the intra refresh line were dropped (Drop Below IR). For the last test, and the one used as a benchmark for comparison, packets were dropped randomly from all regions (Drop Any). Figure 5 shows the test results for *Paris* test sequence. Each point is the average video quality assessed over 20 runs. The error bars represent the standard deviations from the mean. Figure 5 shows that dropping packets from region 1 (Above IR) results into a significant quality drop up to 2 dB compared with the benchmark of dropping randomly. This is due to the fact that region 1 was recently refreshed

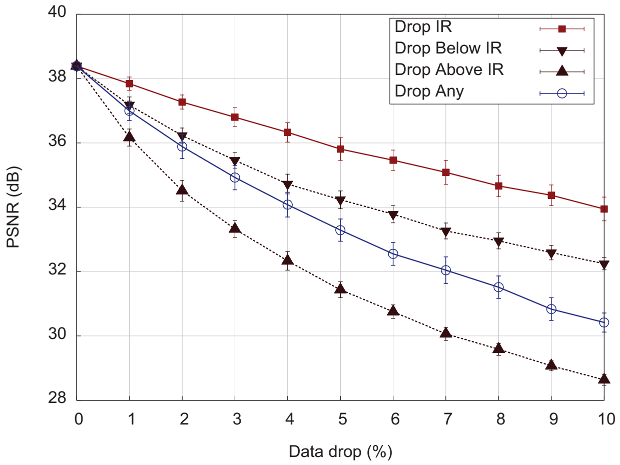

Fig. 5. Quality impact of dropping different regions of the frame for Paris

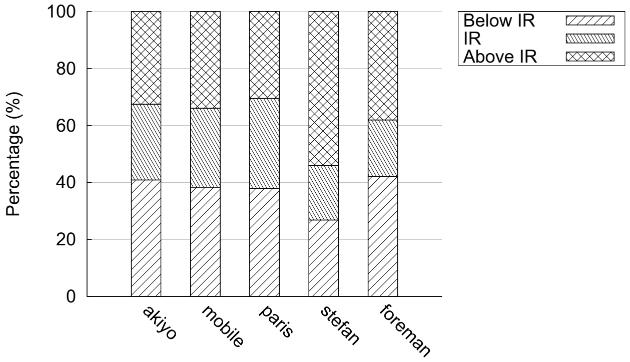

Fig. 6. Percentage of bitrate share for each region

with intra coded blocks and therefore any errors occurring in this region are to be propagated for several frames until these blocks are intra coded again. Dropping packets from region 2 (Drop IR) shows the least impact on video quality (4 dB better than the benchmark). Although intra refresh lines play an important role at mitigating error propagation, these intra coded blocks are very costly (∼20%) as it can be seen in figure 6, while yet representing only a small portion of the image area (5% for a CIF sequence). Dropping packets from region 3 (Drop Below IR) in figure 5 shows a quality gain around 2 dB against the benchmark. Similar results were obtained for different sequences as shown in table I. Overall we conclude that when intra refresh lines are used as an error mitigation scheme it is better to split the frame into three regions as described and implement an unequal protection scheme for each region giving the highest protection to region 1 and the lowest protection to region 2.

V. Conclusion

This paper presented a novel slicing technique for enhanced error robustness. It was found that when intra refresh lines are used as an error mitigation tool, unequal error sensitivity appears within the frame. This fact is described in this paper and a new slicing technique exploiting this characteristic was hereby presented. The proposed slicing scheme was compared with other slicing techniques and found to offer an increased

TABLE I
Results

Sequences	Scheme	Quality Gain versus Drop Any (dB)				
		2%	4%	6%	8%	10%
Akiyo	Above	-1.6	-2.2	-2.5	-2.5	-2.6
	IR	1.0	1.5	1.9	2.2	2.5
	Below	0.9	1.5	2.1	2.6	2.9
Bus	Above	-1.9	-2.5	-1.9	-1.7	-1.7
	IR	2.8	3.3	4.1	4.0	4.0
	Below	1.1	1.5	1.9	2.2	2.6
Foreman	Above	-0.6	-1.1	-1.0	-0.8	-0.8
	IR	2.3	2.7	3.2	2.9	2.8
	Below	0.1	0.1	0.5	0.7	1.0
Mobile	Above	-1.0	-1.7	-1.9	-2.0	-2.0
	IR	1.6	2.6	3.3	3.8	4.1
	Below	0.3	0.6	0.9	1.2	1.4
Paris	Above	-1.4	-1.7	-1.8	-1.9	-1.8
	IR	1.4	2.2	2.9	3.1	3.5
	Below	0.3	0.6	1.2	1.4	1.8
Stefan	Above	-0.4	-0.5	-0.5	-0.6	-0.5
	IR	1.4	2.5	3.0	3.3	3.3
	Below	0.2	0.7	1.1	1.5	1.9

video quality up to 4 dB. Unlike other solutions found in the past, the current implementation does not introduce any computational or bitrate increase, being therefore specially suitable for mobile applications like video calls and DVB-H.

The slicing scheme proposed in this paper adds the possibility to offer unequal error protection using unequal Forward Error Correction codes, quality of service like 802.11e or simple packet duplication.

References

[1] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, "Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T rec. H.264/ISO/IEC 14 496-10 AVC," JVT-G050 2003.

[2] S. Wenger, "H.264/AVC over IP," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, pp. 645 – 656, 2003.

[3] Q. Wang, R. Hu, Z. Wang, and B. Hang, "Intra coding and refresh based on video epitomic analysis," in *Multimedia and Expo (ICME), 2010 IEEE International Conference on*, 2010, pp. 452 –455.

[4] C.-M. Chen, C.-W. Lin, and Y.-C. Chen, "Adaptive error-resilience transcoding and fairness grouping for video multicast over wireless networks," in *Communications, 2007. ICC '07. IEEE International Conference on*, 2007, pp. 1661 –1666.

[5] J. Jiang, B. Guo, and W. Mo, "Efficient intra refresh using motion affected region tracking for surveillance video over error prone networks," in *Intelligent Systems Design and Applications, 2008. ISDA '08. Eighth International Conference on*, vol. 2, 2008, pp. 242 –246.

[6] K. Tan and A. Pearmain, "An FMO based error resilience method in H.264/AVC and its UEP application in DVB-H link layer," in *Multimedia and Expo (ICME), 2010 IEEE International Conference on*, 2010, pp. 214 –219.

[7] J. Xu and Z. Wu, "Joint adaptive intra refreshment and unequally error protection algorithms for robust transmission of H.264/AVC video," in *Multimedia and Expo, 2006 IEEE International Conference on*, 2006, pp. 693 –696.

[8] T. Wiegand, G. Sullivan, G. Bjntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, pp. 557–559, July 2003.

[9] S. Moiron, I. Ali, M. Fleury, and M. Ghanbari, "A prioritization scheme for H.264/AVC video congestion resilience with ieee 802.11e," in *Wireless Advanced (WiAD), 2010 6th Conference on*, 2010, pp. 1 –6.

[10] http://iphome.hhi.de/suehring/tml/, *Suehring JM H.264*.