An Unequal Error Protection in ROI for H.264

Jia Ma, Hongkai Xiong, Li Song, Songyu Yu

Institution of Image Communication and Information Processing
Department of Electronics Engineering
Shanghai JiaoTong University, Shanghai, China, 200240

Abstract. Compressed video sequences are very vulnerable to channel disturbances especially when they are transmitted over the packet erasure channels. The transmission errors[1] caused by packet loss not only corrupt the current decoded frame, but also cause the error propagation to the succeeding frames which degrades the quality of services. Therefore, it is considerably important to develop the error resilience techniques to minimize the visual degradation caused by the packet erasure problem. In this paper, we present a ROI (region of interest) protection strategy which aims to reduce the visual distortion caused by packet loss and also obtain a better balance between the transmission overlay and the perceptual benefits. In our method, FEC protection is applied to the ROI in a chessboard-distributed manner and transmitted through the RTP/RTCP Protocols over UDP/IP. Experimental results show that the proposed algorithm has satisfactory subjective and objective video quality in the burst packet-loss environments.

Index Terms—unequal protection, error resilience, ROI, H.264/AVC,

1. INTRODUCTION

Nowadays, with the huge expanding demands of multimedia services such as real time streaming and broadcast/multicast services, certain relative technologies have been developed rapidly including the emergency of more coding efficiency standards to relieve the congestion situation of the networks. However, the larger degree of compression of the original video contents also means the greater vulnerability of the compressed contents to the channel disturbances as they should be more correlated and interactive. Transmission errors caused by the channel disturbances such as packet losses not only corrupt the current decoded frame, but also cause the error propagation to succeeding frames which deteriorates the quality of services. And in some extreme instances, loss of the packets containing the most important information, e.g. the Parameter Set for H.264 bit stream, even causes the failure of the entire decoding process. Therefore, it is considerably important to develop appropriate error resilience techniques to minimize the visual degradation caused by all kinds of channel disturbances, among which the one we inferred here is mainly about the packet loss problem.

Generally, error resilience techniques can be divided into three categories, depending on the role that the encoder, decoder, or the network layer plays in the process [1]. For the encoder side ER (error resilience) techniques, these ER encoders typically are less efficient in that they use more bits to obtain the same video quality in the absence of any transmission errors. But if the bit stream is corrupted by transmission errors, the extra bits can be used to greatly enhance the video quality. And as for the decoder side ones, various error concealment steps are applied to the corrupted bit stream for reconstruction use taking advantage of the available redundancy information or the inherent spatial/temporal correlation between frames. Obviously, better error concealment results will be achieved if there is more redundancy information provided in the transmitted bit stream. Finally, network layer ER techniques are combinations of the two former ones. Feedback information is provided by the decoder to adapt the encoder ER to achieving the maximum gain with the smallest amount of redundancy. Therefore, we can conclude that more bits allocated for redundancy information makes the bit stream more robust but less efficient in all kinds of ER techniques and the target is always to find the optimized tradeoff.

In this sense, the latest video compression standard H.264/AVC remarkably exceeds all the former ones in the coding efficiency, thus ensuring the better implementation of the effective ER techniques such as data partitioning, RS (redundant slices) and FMO (flexible macroblock ordering) for a wide range of video consumer applications [2]. And in this paper, we propose our new protection strategy which applies the FEC protection in a chessboard-distributed manner to the ROI transmitted through the RTP/RTCP Protocols over UDP/IP. The experimental results will be analyzed and discussed in comparison with that of the data partition technique contained in the JM.

This paper is organized as follows. Our new protection method is described in section 2. The experimental results and some discussion of the new scheme are presented in section 3. Finally, section 4 concludes the paper and indicates some future work.

---

[1] This work is supported by the grant of NSF No.60502033 and Shanghai NSF No.04ZR14082.

2. Proposed Unequal Protection Strategy

As we all know, the distortion of a frame usually consists of two parts: the source distortion due to quantization errors and the channel distortion due to loss of data. Although both of them contribute to the degradation of visual quality, the latter one affects more in most cases as it will cause the potential error propagation problem. Unlike the dispersed visual degradation due to higher QP values, the degradation caused by error propagation actually assembles the unpleasing impacts in a more concentrated region of the frame which is much more visually unbearable for human eyes. Therefore, we try to depress the more sensible channel degradation while having to tolerate a slightly increased source distortion as a tradeoff. In other words, some bits will be allocated for the FEC protection instead of the encoded content to improve the visual quality when transmitting over packets-erasure channel.

For description convenience, assume that the encoded ROI data forms *k* packets for transmission use. And our strategy is applied to half of the macroblocks in the ROI which together shape a chessboard and will be referred to as the “protected MBs” hereafter. To generate the redundant FEC packets, we should gather all the “protected MBs” first to form another *k’* “modified packets”. For example, if the ordinary raster scan order is used during encoding, we will intentionally excluded all the data of odd MBs in even rows from the packets and vise versa to make the *k’* modified packets. In such a case, *k’* is equal to *k* but the packet size is approximately halved. Also, if FMO chessboard map is used, we can simply pick the *k’* data packets that contain slices belonging to the given slice group from the whole *k* packets and *k’* is just half of *k* now. In general, the “gathering” step is implemented by referring to the slice group map and *k’* is no greater than k in all the cases. After gathering, the *k’* modified packets will be padded individually to ensure that each packet has the same length, which is determined by finding the maximum packet length among the *k’* modified packets. And then these new packets will be FEC coded resulting *n* redundant FEC packets where *n* is decided with the consideration of the current network status. The n FEC packets will be transmitted after *k* data packets. And then by correctly receiving/generating at least *k’* FEC/“modified packets” at the decoder, the *k’* modified packets containing half of the MBs in the ROI will be recovered successfully.

One thing to mention here is the reason that we choose such a chessboard-distributed-map when protecting. We do so in order to reduce the FEC redundancy while taking the error concealment step into account at the same time. When packet loss occurred during transmission, error concealment should be employed at the decoder to improve the perceptual effect. In fact, with the adoption of some effective error concealment algorithms, if the spatially adjacent MBs around a lost one are received correctly, the visual impact of that loss can be negligible in the rough [3]. For example, in [4], an error concealment algorithm for H.264 standard compressed video is presented which makes the selection of the lost block type using information provided by the available surrounding blocks. And in [5][6], neighboring MBs also help a lot for the estimation of the lost MB motion vectors. Therefore, the chessboard distributed MBs we choose to protect not only enhance their own correct reconstruction but also offer sufficient prior information for error-concealment step if the “unprotected MBs” are lost.

The protection weight for the ROI data is decided cautiously by the network conditions as mentioned above to avoid overprotection which will consume useful bandwidths on one hand, and on the other hand to prevent under-protection which will result in the degradation of visual quality. To give a concrete discussion, we simply assume that each frame of the sequence is encoded in raster scan order and each row of the MBs is encapsulated in a single packet. That is to say, the number of the modified packets *k’* is equal to that of the data packets k as is discussed above.

If we simply consider a packet erasure channel where packets are either lost or correctly received and the loss ratio of each packet is equal and denoted by p. Also, we denote by $P_{k+n}(i)$ the probability of receiving exactly *i* packets of *k*+*n*. And let *P* denote the probability of the status that less than *k* packets out of *k*+*n* are correctly received and thus fail to be recovered as formulated below.

$$P = \sum_{i=0}^{k-1} P_{k+n}(i) \text{, where } P_{k+n}(i) = C_{k+n}^{i} (1-p)^{i} \cdot p^{(k+n-i)} \qquad (1)$$

Table1 lists the different outputs of *P* while packet loss ratio *p* varies from 5% to 20% and the number of FEC packets *n* varies from 1 to 6. And *k* is fixed to 6. Also, the approximate redundancy ratios caused by different numbers of FEC packets are listed alongside as another important factor for the decision of the *n* value.

However, the real network conditions are far more complicated than the simplified characteristics of the binary symmetry channels. When transmitting over the real lossy channels, most of the losses appear as successive packets losses called “burst losses” instead of the isolated packet losses. Besides, the length of a burst loss is shown to have an important effect on the resulting distortion, where longer burst lengths generally led to larger distortions [7]. Therefore, it is practical to take into account the burst losses.

In [8], a sine model is shown to have well explained the characteristics of the burst packet losses characteristics as measured. Simulation results also showed that when the packet-loss ratio was high (R = 0.41), longer burst packet losses occurred and when the packet-loss ratio was low (R = 0.04), the burst length was limited to about ten. But generally speaking, about 80 to 90% of all burst losses had a length of less than four. Hence, we further set the minimum number of the FEC packets to 3 to facilitate a simplified adaptation.

Table 1. Probability of incapability of recovering.

P		Packet Loss Ratio *p*				
		5%	10%	15%	20%	
n	1	4.44	14.97	28.34	42.33	5.56
	2	0.58	3.81	10.52	20.31	11.11
	3	0.06	0.83	3.39	8.56	16.67
	4	0.01	0.16	0.99	3.28	22.22
	5	0.00	0.03	0.27	1.17	27.78
	6	0.00	0.01	0.07	0.39	33.33

3. Experimental Results and Discussion

In our experiments, we use the H.264 reference software JM10.1 as the codec and the error concealment module is contained in the post-processing step of the decoder. The burst packet lossy nature of the IP network is simulated by discarding some IP packets depending on the information available in a supplied error pattern file. The error pattern file we used in our experiments is the one with average packet loss rates of around 10%. See [9] for details on their characteristics and how they were generated.. After considering both the network condition and the stream overload as mentioned above, we choose 1% as the threshold of *P* (i.e., 1-*P*=99%), which means that every six data packets will eventually form 3 additional FEC packets.

Besides, to give a clear comparison, another error resilience tool, i.e. the data partition technique, is applied to the same content. The data partition technique is proved to greatly enhance the error resilient ability of the transmitted bitstream except that it relies completely on the superior error concealment mechanisms that are available when it can be made sure (or almost sure) that at least the partition A arrives[]. In fact, loss of even a single packet containing partition A data will cause the holdup of the decoder which will further lead to an unbearable drop of the visual quality until the arrival of the next I frame. Therefore, excessive protection

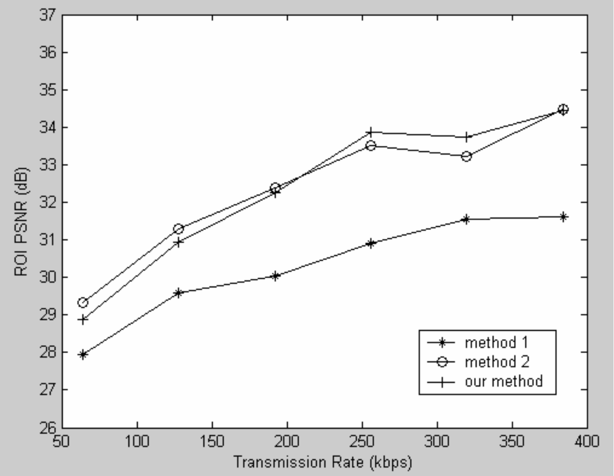

Fig. 1. Comparison of RSNR in the ROI.


(a)
(b)
(c)

Fig. 2. Comparison of the 13th decoded frame of Foreman sequence encoded with (a) method 1: slice, (b) method 2: data partition, (c) method 3: our method.

should be applied to all of the packets containing partition A data, which can be implemented through different kinds of methods. In our experiments, we simply transmit all the packets containing partition A data twice and suppose that all of them arrive at the decoder with no loss.

Figure 1 shows the comparison of PSNR in the ROI when the Foreman QCIF sequence is encoded with method 1: slice, method 2: data partition and our method, respectively. Rate control is employed at the encoder side to generate various bit rates. We convert the bit rate in advance in order to keep the total bit rate in all of the approaches approximately the same. And the subjective comparison is also given in Figure 2.

Obviously, the result of our method is better than that of method 1 and comparable to that of method 2. Subjective comparison showes that both our method and method 2 give a better reconstruction of the image than method 1 does. Besides, method 2 outperforms our method in the reconstruction effect of the texture information (roof and edge of the hat), but is not as good as our method in the representation of the details (wrinkles and the watermark). This is because heavier protection was applied in method 2 to make sure that all of the packets containing the most important data, i.e. the partition A, would be successfully accepted and hence led to a better error concealment result. However, more bits allocated for protection use also meant the decrease of the available bits for encoded video content and thus resulted in coarser QP values and loss of details. Also, without strict requirements on successfully receiving specific packets, our method seems more robust with this understanding.

4. Conclusion

To conclude, by applying protection to chessboard-distributed-MBs in the ROI, these “protected MBs” are more likely to be received/recovered correctly and thus lead to a decrease of the potential error propagation due to both packets loss and dissatisfied error concealment. However, the visual benefits from the additional protection means less available bits for encoded data at the same time, which will result in coarser value of QP and more source distortion if the total amount of bits is restricted. Therefore, a reasonable bits allocation between the encoded video data and the protection data will better adapt the encoded content to the packets-erasure channel and optimize the available perceptual effects. Moreover, rational order of the packets sent may help a lot in conquering the burst error problem and therefore decrease the protection weight and the redundant overload which can be further investigated in the future work.

References

1. Yao Wang, Wenger, S., Jiantao Wen, Katsaggelos, A.K.: Signal Processing Magazine, IEEE. Error resilient video coding techniques. Volume 17 of Issue 4 (2000) 61 - 82
2. Stephan Wenger: IEEE transaction on circuits and systems for video technology. H.264/AVC Over IP. Volume 13 of No.7 (2003) 645-656
3. Yao Wang, Qin-Fan Zhu.: Proceedings of the IEEE. Error control and concealment for video communication: a review. Volume 86 of Issue 5 (1998) 974 – 997
4. Pei-Jun Lee, Chen, H.H, Liang-Gee Chen.: Intelligent Multimedia, Video and Speech Processing. A new error concealment algorithm for H.264 video transmission. (2004) 619 – 622
5. Jinghong Zheng, Lap-Pui Chau. : Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference. An efficient spatial domain error concealment method for H.264 video. Volume 1 (2003) 26-30
6. Jinghong Zheng, Lap-Pui Chau. : Multimedia, IEEE Transactions. Efficient motion vector recovery algorithm for H.264 based on a polynomial model. Volume 7 (2005) 507-513
7. Liang, Y.J, Apostolopoulos, J.G.; Girod, B. : Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference. Analysis of packet loss for compressed video: does burst-length matter? Volume 5 (2003) V - 684-7
8. Arai, M.; Chiba, A.; Iwasaki, K. : Dependable Computing, 1999. Proceedings. 1999 Pacific Rim International Symposium. Measurement and modeling of burst packet losses in Internet end-to-end communications. (1999) 260 – 267
9. [38] VCEG. (1999) Internet Error Patterns VCEG-O38r1.doc.