ERROR-RESILIENT WIRELESS VIDEO TRANSMISSION USING MOTION-BASED UNEQUAL ERROR PROTECTION AND INTRA-FRAME PACKET INTERLEAVING

Qi Qu †*, *Yong Pei*‡ , *James W. Modestino*† *and Xusheng Tian*†
† Electrical and Computer Engineering Dept., University of Miami, Coral Gables, FL33124, USA
‡Computer Science and Engineering Dept., Wright State University, Dayton, OH45435, USA

ABSTRACT

Packet video transmission over wireless networks is expected to experience packet losses due to noise, interference, multipath and the motion of mobile hosts. Such packet losses often occur in bursts which may cause substantial degradation to the transmitted video quality. In this paper, we propose the use of a cross-layer unequal error protection (UEP) approach which is achieved by assigning an unequal amount of forward error correction (FEC) to each group of link-layer packets according to the corresponding motion level of the slice acquired at the application layer. Also, a novel packetization scheme and an intra-frame packet interleaving scheme are proposed to be used together with FEC/UEP in order to combat the bursty packet losses in 3G wireless environments. Our results demonstrate that the proposed approach is very effective in dealing with the packet losses occurring on wireless networks without incurring any additional implementation complexity or delay.

Keywords: Video transmission, motion level, FEC/UEP, intra-frame interleaving, bursty packet loss.

1. INTRODUCTION

With the advent of third-generation (3G) wireless cellular systems, the transport of multimedia, such as video and audio, over wireless networks has attracted increased attention. However, the characteristics of wireless systems provide a major challenge for reliable transport of multimedia applications since the data transmitted over wireless channels are highly sensitive to the noise, interference, and multipath environment which can cause both packet losses and bit-errors. Furthermore, these errors tend to occur in bursts and, as shown in [1], burst packet losses will generally cause much larger total distortion than random packet losses for the same average packet loss rate.

In wireless networks, channel feedback is hard to obtain in a timely and reliable way due to the rapid change of wireless environments. However, source information is always available and can be obtained easily and reliably. Therefore, in this paper we propose error-resilient techniques for video delivery over wireless networks based on source adaptation alone.

More specifically, in this work we explore a cross-layer FEC/UEP scheme based on the motion information of transmitted video. For a video frame, some portions have more inter-frame motion than others. The loss of these high-motion portions can cause relatively larger distortion compared to other portions due to the increased perceptual importance of this high-motion information as shown in [4]. Clearly, we need to protect the high-motion portion with higher FEC protection, while lower FEC protection should suffice for the less significant low-motion portions. In this paper, we consider an H.264 encoder/decoder and take the level of motion associated with a slice as indicative of the relative importance of the corresponding data. A slice is composed of an integer number of macroblocks (MBs). We assume YUV image data in QCIF 4:2:0 format. The motion levels associated with a slice are classified in terms of the mean-square values of the corresponding inter-frame prediction errors. We then use different Reed-Solomon (RS) codes to protect the slice depending on the computed inter-frame motion levels thereby achieving UEP.

In this work, we employ a novel packetization scheme, based on the UMTS protocol architecture, which can simultaneously provide efficient source coding performance and robust delivery. Furthermore, in order to randomize the bursty packet losses, we employ an intra-frame interleaving scheme which interleaves the slices (the application packets) within a frame. Thus, since the delay is constrained within a single video frame, no additional delay is incurred, while this scheme is still capable of significantly randomizing the burst losses occurring on wireless networks. Therefore, improved performance can be expected.

The rest of the paper is organized as follows: in Section 2, we discuss the proposed packetization scheme; in Section 3, the cross-layer motion-based FEC/UEP approach is described; the intra-frame interleaving scheme is discussed in Section 4; selected simulation results are presented in Section 5 followed by conclusions in Section 6.

2. PROPOSED PACKETIZATION SCHEME

The introduction of slices in the encoding process has at least two beneficial aspects for video transmission/streaming over wireless networks. The two primary factors are the reduced error probability of shorter packets and the ability to resynchronize within a frame [2]. However, this also adversely affects the source coding efficiency due to the increased slice overhead and reduced prediction accuracy within a single frame, since inter-frame motion vector prediction and intra-frame spatial prediction are not allowed across slice boundaries in H.264 [2]. In our previous work [1], and also in [2], it is demonstrated that 6-9 slices per QCIF frame is a reasonable choice for a wide range of operating bit rates and channel conditions

When transmitting over the wireless network, the application layer video packets are segmented into radio link protocol (RLP) packets at the link layer. This segmentation can cause some problems. For example, if the existing transport protocol is UDP, unless all the RLP packets are received successfully, the entire application packet will also be lost unless the UDP error checking mechanism is disabled. Since UDP has other desirable properties compared to TCP for real-time video transport applications, we will concentrate on the use of UDP in this paper.

Based on the discussion above, our proposed packetization approach is implemented as follows:

---

*Qi Qu is now with the ECE Dept., University of California San Diego, La Jolla, CA92093.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

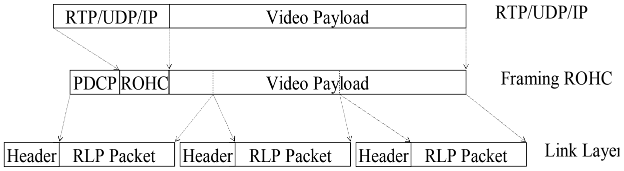

Fig. 1. Proposed packetization for UMTS wireless networks.

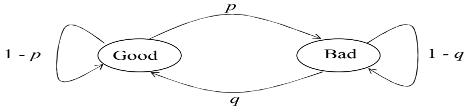

Fig. 2. Two-State Gilbert Model.

1. In the encoding process, each slice in a frame consists of an equal number of MBs (in this paper, we exclusively use 11 MBs per slice, thus every QCIF video frame is divided into 9 slices). Then, each encoded slice is packetized into one RTP/UDP/IP packet, also called an application packet.
2. Since the induced packet overhead for every RTP/UDP/IP packet is 40 bytes, in order to economize on the scarce bandwidth resource, we use robust header compression [3] to compress the RTP/UDP/IP header into 3 bytes.
3. At the link layer, each application packet is divided into several equal-size RLP packets, say M, according to the MTU size of this transmission. The value of M is kept constant for the whole transmission session. We add a header to every RLP packet, one part of which can be used to allow the FEC decoder to determine the positions of lost packets and the other part can be used to indicate which FEC code is used for the FEC/UEP scheme as will be discussed later.
4. Then, the proposed FEC/UEP scheme in Section III is applied to the set of RLP packets which belong to the same application packet. The data packets together with the parity packets are then delivered over the network.
5. Finally, at the receiver, the FEC decoder first recovers the lost packets and if every RLP packet within an application packet is received correctly, the corresponding application packet is delivered to the upper layer; if not, the corresponding application packet is discarded.

The process is indicated in Fig. 1, for example, for 3G UMTS wireless networks. We model the packet loss for video delivery at the link-layer level. The loss model we use is the two-state Gilbert Model indicated in Fig. 2, which can be uniquely specified by the average burst length (L_B) and the average packet loss rate (P_L). They can be explicitly related to the corresponding state transition probabilities p and q as described in [1].

3. A CROSS-LAYER UEP APPROACH

In previous work [4], motion has been determined by a statistical classifier; namely, by the inter-frame prediction error which is used as a primary indicator of activity of video frames. We will also use this statistical classifier in our motion-based adaptive system to classify the motion levels of slices.

For the transmitted video sequence, at the application layer we first calculate the prediction error between slices which are in the same positions in successive frames according to

$$E[m,n]=\sum_{j=0}^{N_v-1}\sum_{i=0}^{N_h-1}\left[X_{m,n}(i,j)-X_{m-1,n}(i,j)\right]^2, \qquad (1)$$

where $E[m,n]$ denotes the mean-square prediction error between the luminance data in the n-th slice of the m-th frame, of size N_v x N_h pixels, and the corresponding data in the n-th slice of the (m-1)-th frame of the video sequence where the total number of frames is N_f.[1]

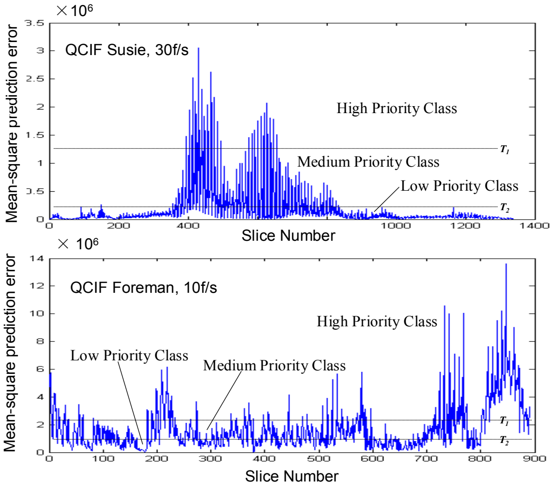

Fig. 3. Slice prediction error for QCIF Foreman and Susie.

In Fig. 3, the slice prediction errors are indicated for the QCIF Foreman and the QCIF Susie sequences at 10 fps and 30 fps, respectively. The results are consistent with subjective observations for these two sequences, where the Susie sequence is considered to have low motion with the background tending to remain constant. On the other hand, the Foreman sequence has much more motion due to activity and scene changes. Also, for each of them, some portions of frames have more inter-frame motion than others. Our FEC/UEP scheme is based on the slice prediction errors acquired at the application layer. As noted in Fig. 3, we set two thresholds T_1 and T_2, illustrated in both plots, which are chosen as indicated later. We classify the slices in a video sequence according to the motion levels:

High-Priority Class: slice prediction error above T_1;
Medium-Priority Class: slice prediction error within T_1 and T_2;
Low-Priority Class: slice prediction error below T_2;

In our approach, UEP is realized by assigning an unequal amount of FEC at the link layer to classes with different motion levels. This is a cross-layer approach since the link layer requires the motion information obtained at the application layer to implement the FEC/UEP scheme. For each slice in the same class, the same RS code is applied across the RLP packets after the corresponding application packet is split into M equal size RLP packets, i.e., for M RLP packets in the same application packet, a RS(n, M) code is applied. At the receiver side, after the M RLP packets for a single application packet are received, the FEC decoder can identify the positions of lost packets using the header information of each RLP packet. The lost packets can be recovered if the number of lost packets is less than $n - M$. If any one of the RLP packets cannot be recovered, the entire application packet is discarded.

Thus, since different classes use an unequal amount of redundancy, UEP is achieved. By using this approach, slices with higher motion are protected by stronger RS codes, while slices with lower motion are protected by weaker RS codes. In this way, the system bandwidth resources can be utilized more efficiently. Actually, the loss of a low-motion frame/slice is barely noticeable and can be effectively concealed by the built-in passive error concealment together with intra-updating. However,

---

[1] In this paper, each video frame has the same number of slices which are the same size and in the same position.

the loss of a high-motion frame/slice may cause substantial performance degradation in video quality due to severe error propagation effects.

We should also note that in this approach, the FEC coding is applied across the RLP packets within a single application packet. As a result, there is no noticeable delay introduced by this FEC/UEP approach. That is because at the encoder side, after the application packets within one frame are split into several RLP packets, we can buffer the RLP packets while simultaneously sending the RLP packets to the receiver. The buffered RLP packets can then be used in the FEC encoding process to compute the parity packets so that no need exists to delay the transmission of the RLP information packets. Likewise, at the receiver side, a decoding delay is incurred only for those application packets with lost RLP packets, otherwise no delay is induced. When implementing this FEC/UEP system, since different FEC codes are used for any given slice, the receiver requires information about which RS code is applied. Obviously, the receiver can be notified of this through a field in the RLP packet header.

4. INTRA-FRAME INTERLEAVING

Since in wireless networks, packet losses tend to occur in bursts, which can cause a serious performance degradation, it is desirable to randomize the burst packet losses. For traditional interleaving schemes, this is generally achieved by interleaving the application packets over several successive frames; thus, a large delay is introduced. This is not practical for real-time video applications.

Therefore, in this paper, we propose an intra-frame interleaving scheme which only interleaves the slices within a single frame. Thus, no additional delay is introduced for real-time video transmission. In cases where two or more successive slices within a frame are lost due to the burst errors on wireless channels, this interleaving scheme can randomize the resulting bursty losses at the slice level, thus rendering the built-in passive error concealment (PEC) more effective [1]. Specifically, the built-in motion-compensated PEC algorithm [6] proceeds by first examining the motion activity of the correctly received slices of the current frame; if the average MV is smaller than a pre-defined threshold, all lost slices are concealed by copying from the spatially corresponding positions in the reference frame. Otherwise, motion-compensated error concealment is used: the motion of a "Lost" MB is predicted/recovered from spatially neighboring MB's relying on the statistical expectation, that the motion of spatially neighboring areas is highly correlated with that of the “Lost” MB. Thus, successful reception of the neighboring slices is important for high performance recovery of any single lost slice. In Fig. 4, we illustrate the proposed application-layer interleaving scheme. Since in this paper each frame is split into nine slices, we interleave the positions of the nine slices within the frame according to the pattern illustrated in Fig. 4. As can be seen from Fig. 4, if we assume no interleaving and a burst loss of length 3 affecting slice#0, slice#1 and slice#2, then at the decoder side it is difficult for the motion-based PEC scheme to conceal the effects of this burst loss since no neighboring MBs are available for slice#0 and slice#1, and for slice#2 only slice#3 is available. Therefore, the performance of the PEC scheme is degraded significantly. However, if application-layer interleaving is applied, then for the same burst loss, after interleaving and subsequent deinterleaving, the burst loss is randomized to some extent as illustrated in Fig. 4. In this case the performance of the PEC scheme can be significantly improved since the necessary information for operation of the PEC algorithm is available from neighboring MBs.

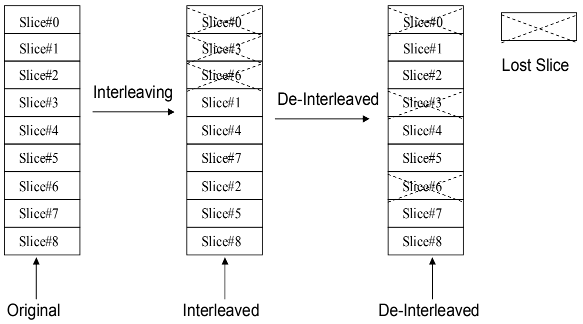

Fig. 4. Intra-frame interleaving scheme.

5. SIMULATION RESULTS AND DISCUSSIONS

We provide simulation results to demonstrate the potential performance gain that can be achieved by the proposed cross-layer FEC/UEP approach together with the proposed packetization and intra-frame interleaving scheme over 3G UMTS wireless networks. Video sequences are encoded using the ITU-JVT JM3.9 codec of the H.264 video coding standard. We will use two typical QCIF test video sequences: Foreman and Susie, as described previously. Both are coded at constant bit rates specified by using the associated H.264 rate control scheme. The first frame of the sequence is intra-coded and the rest of the frames are inter-coded as *P* frames with intra-updating of one MB line (slice) each frame. For the motion-based FEC/UEP scheme, a RS(6, 3) code is used for the high-priority class, a RS(5, 3) code is used for the medium-priority class and a RS(4, 3) code is used for the low-priority class. For comparison, we also use an equal error protection (EEP) scheme without interleaving; here the packetization process is the same and we employ a RS(5, 3) code for all classes. The thresholds T_1 and T_2 indicated previously are then chosen so that the overall channel coding rates of these two systems are approximately equal. Also the link-layer retransmission function is disabled.

In Fig. 5, we show the results for the Foreman sequence for burst length L_B = 3, where we choose the thresholds such that out of a total of 900 application packets, 205 application packets are classified into the high-priority class, 500 packets into the medium-priority class and 195 packets into the low-priority class. The average channel coding rate is then 0.608 bits/c.u compared to 0.60 bits/c.u for the EEP scheme. In Fig. 5, by using the FEC/UEP scheme, lower-priority classes are provided with lower protection, while higher-priority classes are provided with higher protection since the packet losses of the low-priority class will contribute smaller total distortion compared to the packet losses of the high-priority class. Thus, an efficient way of distributing the FEC redundancy over different classes can be achieved. For the burst-loss channel, the use of our proposed scheme can also randomize the burst errors to some degree due to the use of intra-frame interleaving.

Therefore, the proposed UEP scheme should outperform the EEP scheme, especially for high packet loss rate. For both R_{tot} = 96 Kbps and R_{tot} = 256 Kbps, when the packet loss rate is higher than 5%, a performance gain of approximately 2 dB is achieved by the proposed UEP scheme compared to the EEP scheme.

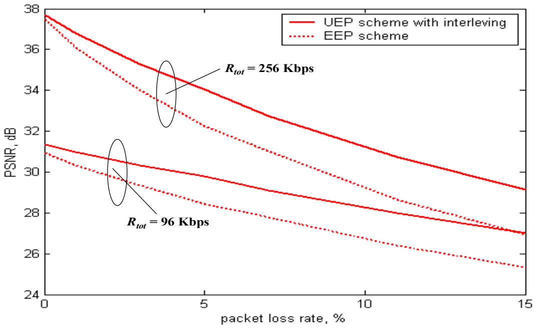

Fig. 5. Proposed UEP scheme *vs*. EEP scheme; L_B = 3; the Foreman Sequence.

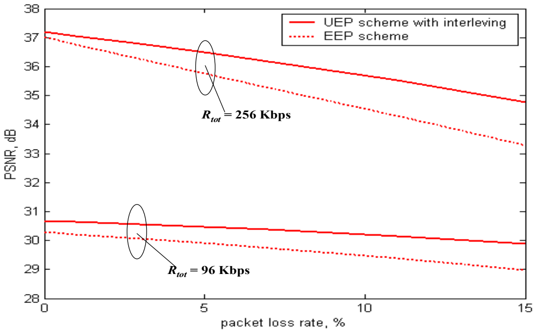

Fig. 6. Proposed UEP scheme *vs*. EEP scheme; L_B = 3; the Susie Sequence.

While in the lossless case, the UEP scheme can achieve a performance gain of about 0.5 dB.

In order to provide a fair evaluation of our proposed scheme, we repeat the simulations for the QCIF Susie sequence which has a much lower overall motion level than the Foreman sequence. The results are illustrated in Fig. 6, again for $L_B = 3$.

For the results illustrated in Fig. 6 for the low-motion Susie sequence, we observe that the proposed UEP scheme with intra-frame interleaving still achieves a higher performance than the EEP scheme. For example, at $R_{tot} = 96$ Kbps, for packet loss rate in excess of 5%, the gain is about 1 dB, while for the lossless case, the gain is about 0.4 dB. For $R_{tot} = 256$ Kbps, when the packet loss rate is 15%, the gain is about 2 dB. Again, the results demonstrate the effectiveness of our proposed scheme compared to the EEP scheme. However, we should also note that for the low-motion Susie sequence, the gain achieved by the proposed scheme is somewhat less than for the moderate-motion Foreman sequence. The reason is that since the overall motion level of Susie is much lower than that of Foreman, as indicated in Fig. 3, the built-in PEC is more effective in dealing with the packet errors in the Susie sequence.

In the preceding results, the average burst length L_B equals 3 for the link-layer packets, which means a corresponding approximately random loss pattern since each application packet is segmented into three equal-size link-layer packets. As demonstrated in [5], the burst length in representative UMTS channels is likely in the range of 1-3 for application packets with high probability. So, the average burst length at the link-layer should be approximately 3 – 9 for the packetization scheme used in this paper. As expected, the simulation results indicate that as the average burst length increases, the effectiveness of this approach decreases to some degree depending on L_B. In Table I, we summarize these results. From the table, we can see that as the burst length increases, the performance gain in terms of PSNR (dB) achieved by the FEC/UEP scheme together with the intra-frame interleaving decreases. The reason is that as the average burst length increases, the intra-frame interleaving scheme becomes less effective.

	$L_B = 6$		$L_B = 9$	
	UEP	EEP	UEP	EEP
P_L=1%	30.74	30.24	30.66	30.12
P_L=5%	28.39	27.88	28.15	27.76
P_L=11%	26.34	25.56	25.76	25.36
P_L=15%	25.13	24.41	24.59	24.20

Table I. Performance comparison of UEP with Interleaving vs. EEP without interleaving, R_{tot} = 96Kbps, the Foreman sequence.

6. CONCLUSIONS

In this paper, we investigated an FEC/UEP scheme based on the inherent motion information of video sources. The video data is classified into three priority classes, each of which is assigned an unequal amount of FEC redundancy in order to provide the appropriate protection to different priority classes. The simulation results demonstrate that this proposed FEC/UEP scheme, used together with the proposed packetization and intra-frame interleaving schemes, can provide improved performance compared to the equal error protection (EEP) scheme. The gain is about 0.3 – 2 dB according to the channel conditions. As the average burst length increases, the performance gain will also decrease to some degree due to the loss of effectiveness of the proposed intra-frame interleaving scheme. Another way to further improve the system performance is to exploit a link-layer interleaving scheme which also operates within a single frame and can randomize the burst errors at the link-layer level. This is being investigated in ongoing work.

7. REFERENCES

[1] Q. Qu, Y. Pei and J. W. Modestino, "Robust H.264 Video Coding and Transmission over Bursty Packet Loss Wireless Networks", in *Proc. of IEEE VTC'2003*, Oct. 2003.

[2] T. Stockhammer, et. al., "H.264/AVC in Wireless Environment", *IEEE Trans. on CSVT*, pp. 657-673 July 2003.

[3] S. Casner, V. Jacobson, "Compressing IP/UDP/RTP Headers for Low-Speed Serial Links", RFC2508, Feb.1999.

[4] M. Bystrom, V. Parthasarthy, and J. W. Modestino, "Hybrid Error Concealment Schemes for Broadcast Video Transmission over ATM Networks", *IEEE Trans. on CSVT*, vol. 9, no. 6, pp. 868-881, Sept. 1999.

[5] S. Gnavi, et. al., "Comparison of Rate Allocation Strategies for H.264 Video Transmission over Wireless Lossy Correlated Networks", *ICME'03*, pp. 517-520, Jul. 2003.

[6] Y.-K. Wang, et. al., "The Error Concealment Feature in the H.26L Test Model", *ICIP'02*, pp. 729-732, Sep. 2002.