INTERACTIVE VIDEO CODING AND TRANSMISSION OVER WIRED-TO-WIRELESS IP NETWORKS USING AN EDGE PROXY

Yong Pei and James W. Modestino

Electrical and Computer Engineering Department
University of Miami
Coral Gables, FL 33124, USA

ABSTRACT

In this paper, we address the problem of enabling robust interactive video coding and transmission over heterogeneous wired-to-wireless IP networks. We propose the use of an FEC coding scheme employing Reed-Solomon (RS) codes and rate-compatible punctured convolutional (RCPC) codes to protect the video data from packet loss and bit errors, respectively. Furthermore, we apply an end-to-end architecture using an edge proxy in a mobile support station to implement differential error protection for the corresponding channel impairments expected on the two networks. Results indicate that with an appropriate joint source-channel coding approach and the use of an edge proxy, FEC-based error-control techniques together with passive error-recovery techniques can significantly improve the effective video throughput and lead to acceptable video delivery quality over time-varying heterogeneous wired-to-wireless IP networks.

Keywords: Video transmission, RTP, UDP/IP, RS Codes, RCPC codes, joint source-channel coding, H.263+.

1. INTRODUCTION

Many multimedia applications will require video transmission over links with a wireless last-hop. However, many existing wired and/or wireless networks cannot provide guaranteed QoS, either because of congestion, or because temporally high bit error rates cannot be avoided during fading periods. Channel-induced losses, including packet losses due to congestion over wired networks as well as packet losses and/or bit errors due to transmission errors on a wireless network, require customized error resilience and channel coding strategies that add redundancy to the coded video stream at the expense of reduced source coding efficiency or effective source coding rates, resulting in compromised video quality.

This work represents an extension of previous work in [1] and [2]. In particular, in [1] we described an approach using edge proxies which did not address the unique FEC requirements on the wired networks. This was followed by work reported in [2] where a concatenated channel coding approach was employed, but without an edge proxy, which attempted to address the distinct FEC requirement of both the wired and wireless networks. However, this approach is not optimal since the coding overhead required on the wired link must also be carried on the wireless link.

In this paper we present a framework for an end-to-end solution for packet video over heterogeneous wired-to-wireless networks using an edge proxy. A JSCC approach is used with an FEC coding scheme employing RS block codes and RCPC codes to actively protect the video data from the different channel-induced impairments

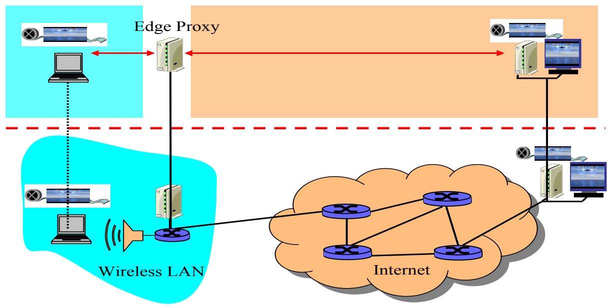

Fig. 1. An End-to-End Approach Using an Edge Proxy.

to be expected over tandem wired and wireless networks.

The remainder of this paper is organized as follows: In Section 2 we present a framework for the proposed end-to-end solution for packet video over heterogeneous wired-to-wireless network using edge proxies. In Section 3 we briefly describe the RS codes and packetization scheme. In Section 4, we present the RCPC codes and channel loss model for the wireless networks under study. In Section 5, we present selected results for the proposed approach. Finally, Section 6 provides a summary and conclusions.

2. PACKET VIDEO OVER WIRED-TO-WIRELESS IP NETWORKS

Video quality should degrade gracefully in the presence of either packet losses due to congestion on the wired network, or bit errors due to fading conditions on the wireless network. Due to the difference in channel conditions and loss patterns between the wired and wireless networks, to be efficient and effective the error-control schemes should be tailored to the specific characteristics of the loss patterns associated with each network. Furthermore, the corresponding error-control schemes for each network should not be designed and implemented separately but jointly in order to optimize the quality of the delivered video.

In this paper, we present an end-to-end solution with the use of a edge proxy operating at the edge of the two network connections as demonstrated in Fig. 1. This end-to-end solution employs the edge proxy to enable the use of different error-control schemes on the wired and wireless networks. Specifically, we employ RS codes on the wired network and RCPC codes on the wireless network to provide error-resilient video service over tandem wired-to-wireless IP networks. As a result, under the constraint of total bitrate budget R_{tot}, the effective video data throughput is given as

$R_s = \min\{R_{tot} \cdot R_c^{RS}, R_{tot} \cdot R_c^{RCPC}\}$, where R_c^{RS} and R_c^{RCPC} are the channel coding rates for the RS and RCPC codes, respectively. In contrast, without the use of an edge proxy, these two codes have to work as a concatenated FEC scheme in order to provide sufficient protection against both congestion-caused packet loss in the wired network and fading-caused bit errors in the wireless network as demonstrated in [2]. Specifically, the RS works as an outer code while the RCPC is used as an inner code. As a result, under the same bitrate constraint, the effective video data throughput without the use of an edge proxy is limited by $R_s = R_{tot} \cdot R_c^{RS} \cdot R_c^{RCPC}$. It is clear that the effective video data throughput can be improved through the use of an edge proxy. In this paper, we will quantitatively investigate the resulting improvement for interactive video coding and transmission with the use of an edge proxy.

For the proposed system, a H.263+ source coder encodes the input video which is applied to a channel encoder employing a RS block encoder and/or a RCPC encoder. The RS code operates in an erasure-decoding mode and provides protection against packet loss due to congestion in the wired IP network while the RCPC provides protection against bit errors due to fading and interference on the wireless network. The RS coding rates can be selected adaptively according to the prevailing network condition; specifically, packet loss rate for the wired IP networks. This channel rate matching is achieved by employing a set of RS codes with different erasure-correcting capabilities. The RCPC coding rates can also be selected adaptively to provide different levels of bit-error correcting capability according to the prevailing wireless network conditions; specifically, E_S/N_I for the wireless channels.

In order to transmit real-time H.263+ video over an IP network, the H.263+ bitstream must first be packetized. The protocol of choice for IP-based real-time packet video applications is the real-time transport protocol (RTP). A payload format for H.263+ video has been defined for use with RTP (RFC 2429) [3]. According to the RTP-H.263+ payload format specification, the H.263+ encoded bitstream is packetized and then transmitted as RTP packets.

Finally, the bitstreams are modulated before being transmitted over a wireless link. During transmission, the modulated bitstreams typically undergo degradation due to additive white Gaussian noise (AWGN) and/or fading. At the receiver side, the received waveforms are demodulated, channel decoded, and then source decoded to form the reconstructed video sequence. The reconstructed sequence may differ from the original sequence due to both source coding errors and possible channel error effects.

In this paper, the symbol transmission rate for the wireless links is set to be $r_S = 64$ Ksps such that the overall bitrate, employing QPSK modulation, is constrained as $R_{tot} = 128$ Kbps. This in turn sets the upper limits for the bit rate over the wired networks to be $R_{tot} = 128$ Kbps as well. Since the total bitrate is limited by the wireless links, as described above, the use of RS and/or RCPC codes will result in a decrease of bitrate used for source coding, i.e., the effective video data rate.

2.1. Edge Proxy

To accommodate the differential error-control schemes as well as differential transport protocols for packet video over wired and wireless networks, appropriate middleware has to be employed to operate between the wired and wireless network to support the application-layer solutions for video applications. Thus we define an edge proxy here to accomplish these functionalities. The edge proxy should be implemented as part of a mobile support station. Furthermore, it should be application-specific; in our case it is video-oriented. The functionalities of the edge proxy include:

1. Selective packet relay
2. Error-control transcoding
3. Joint source-channel coding (JSCC) control
4. Interoperation between different possible transport protocols for the wired and wireless network

For the interactive application we consider here, there exists two-way traffic, including wired-to-wireless as well as wireless-to-wired. Conventionally, block codes are employed to combat packet loss due to congestion in a wired network. Since the channel conditions may vary substantially between the wired and wireless network, such redundant packets in a wired network may not be efficient and effective in the wireless network. As we have shown in previous work [1], constraining the bit error rate to a low level is of primary importance for wireless networks instead of controlling packet loss. It is necessary for the edge proxy to do error-control transcoding if such a scheme is used.

Furthermore, as has been demonstrated in [4], in order to protect against the channel impairments, some form of forward error-control (FEC) coding must be employed. Appropriate FEC coding approaches must be carefully selected and implemented. An important requirement is that they must adequately protect the compressed information bit-stream, according to the characteristics of the given channel, without excessive throttling of the source coding rate to accommodate the coding overheads for time-varying error conditions. Since an arbitrarily chosen FEC design can lead to a prohibitive amount of overhead for highly time-varying error conditions over wireless channels, a joint source and channel coding (JSCC) approach for image or video transmission is necessary. The edge proxy should support the JSCC control scheme to adaptively adjust the source and channel coding rates. To avoid computation- and time- expensive video transcoding in the edge proxy, an end-to-end adaptive coding control strategy is suggested here. The channel conditions, including those for both the wired and wireless networks, are collected in the edge proxy and, based on the prevailing channel conditions, video coding rates are adjusted accordingly using JSCC. For the wired network, the major channel condition parameter is the packet loss rate, while for the wireless channels, channel SNR as well as fading parameters are used.

The edge proxy is also responsible for the interoperation between different possible transport protocols for the wired and wireless network. For a wireless network the error-control scheme is implemented in the application layer, and erroneous packets should be delivered to the end user. However, for conventional wired networks, such as existing IP networks, no error is allowed. In this case, to achieve interoperation, the edge proxy has to repacketize the packet according to the appropriate transport protocol before relaying the packet in either direction.

3. PACKET-LEVEL FEC SCHEME FOR WIRED IP NETWORKS

Packet loss is inevitable, even in the wired IP networks, and can substantially degrade reconstructed video quality, which is annoying for users. Thus, it is desirable that a video stream be robust to packet loss. Regarding the tight delay-constraints for real-time video applications, FEC should be applied to achieve error recovery when packet losses occur. For the wired IP network, packet loss is caused primarily by congestion, and channel coding is typically used at the packet-level to recover from such losses [5]. In

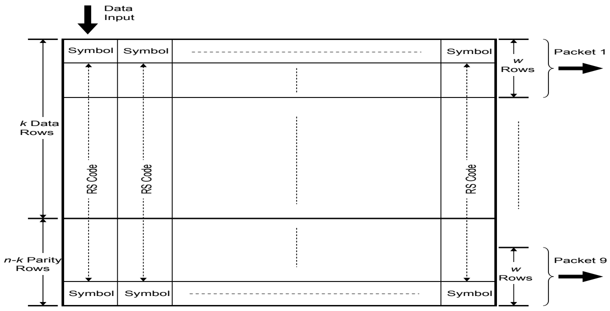

Fig. 2. Illustration of Interlaced Reed-Solomon Codes.

this paper, we will apply a form of interlaced FEC coding employing RS codes, as demonstrated in Fig. 2, where FEC codes are applied across IP packets. Specifically, each packet is partitioned into m-bit symbols and individual symbols are aligned vertically to form RS codewords of block length n over $GF(2^m)$. Then the decoded packet loss probabilities can be readily determined assuming erasure-only decoding [2].

3.1. Packetization for the RS Coded Video Data

To quantitatively compare the performance between the coded system and the uncoded system, we have to maintain the same packet generation rate. Specifically, for the QCIF video studied in this paper, in the uncoded system, each group of blocks (GOB) is packetized into a single packet, resulting in 9 packets per video frame. For the coded system, network packets are obtained by concatenating successive rows of the array illustrated in Fig. 2. We maintain identical packet rate in the coded system as in the uncoded system. Specifically, with the use of RS$(63, k)$ codes, this results in packing 7 coded symbols from the same RS codeword into the same packet together with other RS coded symbols from the same video frame. As a result, both systems will generate 9 packets per frame.

It should be noted that a lost packet in the uncoded system as described above will result in a loss of 1 GOB. However, for the coded system, if there is packet loss that cannot be recovered through the erasure-correcting capability of the corresponding RS codes, the whole frame, i.e., 9 GOBs, will be affected due to the interlaced RS coding scheme. In such a situation, passive error recovery, as will be described in the next section, will be applied to conceal the errors.

4. PACKET VIDEO OVER WIRELESS NETWORKS

4.1. RCPC Channel Codes

The class of FEC codes employed for the wireless IP network in this work is the set of binary rate-compatible punctured convolutional (RCPC) codes described in [6]. With P representing the puncturing period of the code, the rates of the codes that may be generated by puncturing a rate $R_c = 1/n$ mother code are $R_c = P/(P+j),\ j = 1, 2, \ldots, (n-1)P$. Thus, it is easy to obtain a family of codes with unequal error-correcting capabilities. In this work, a set of RCPC codes are obtained by making use of an $R_c = 1/4$ mother code with memory $M = 10$ and a corresponding puncturing period $P = 8$. Then the available RCPC codes are of rates, $R_c = \frac{8}{9}, \frac{8}{10}, \cdots, \frac{8}{32}$.

4.2. Channel-Induced Loss Models

In this work, we restrict our attention to random loss models for both packet loss and bit errors. Specifically, for the wired IP network, packet loss is randomized without consideration of the burst nature of the network congestion. Similarly, the wireless channel is characterized by uncorrelated bit errors, which is a reasonable model for a fairly benign wireless channel under the assumption of sufficient interleaving to randomize the burst errors produced in the channel decoder.

We have shown in [1] the advantage of a transparent transport layer for video transmission over noisy wireless channels. In this paper, we will again assume the transport layer is transparent to the application layer, i.e., a packet with errors is not simply discarded in the transport layer. Instead, the application layer should be able to access the received data although such data may have one or more bit errors. Such a model corresponds to a transport layer scheme allowing bit errors in the payload. The channel-induced impairment to the video quality is then in the form of residual bit errors in the video stream. It is the responsibility of the application layer to deal with the possible bit errors. Specifically, here we make use of the H.263+ coding scheme where, based on syntax violations, certain error patterns may be detected by the video decoder and use of the corresponding errored data can be avoided by employing passive error-recovery techniques.

4.3. Passive Error Recovery

If a packet is considered lost, the RTP sequence number enables the decoder to identify the lost packets, so that locations of the missing data are known. The affected blocks can then be concealed by passive error recovery (PER) techniques. In this work, we make use of the error-detecting and recovery scheme suggested in Test Model 8 [7]. The major objective of this PER scheme is to detect the severe error patterns and prevent the use of such errors which may substantially degrade the video quality. The remaining undetected error patterns in the payload which are not detected by the H.263+ decoder will result in the use of incorrectly decoded image data which can cause quality degradation of the reconstructed video.

5. SELECTED SIMULATION RESULTS

We present some selected results for a representative QCIF videoconferencing sequence, Susie at 7.5 fps. These results were obtained using a single-layer H.263+ coder in conjunction with the proposed FEC coding scheme together with QPSK modulation on the wireless link. A slow and flat Rician fading model for the wireless channel is assumed here. To decrease the sensitivity of our results to the location of bit errors, a sequence of $N_f = 30$ input frames is encoded, channel errors including packet loss and bit error are simulated and the resulting distortion is averaged. Furthermore, each simulation was run N_t times. By taking empirical averages with N_t sufficiently large (i.e., $N_t = 1000$), statistical confidence in the resulting distortion can be achieved.

We first consider the case where no edge proxy is introduced between the wired and wireless networks. In such a case, packet loss due to congestion in the wired network and bit errors due to fading effects in the wireless network coexist. We proposed to jointly select the source coding rate, the RS coding rate and RCPC coding rate in the resulting concatenated scheme such that optimal end-to-end performance can be achieved. Here we demonstrate PSNR results for reconstructed video as a function of the wireless channel E_S/N_I

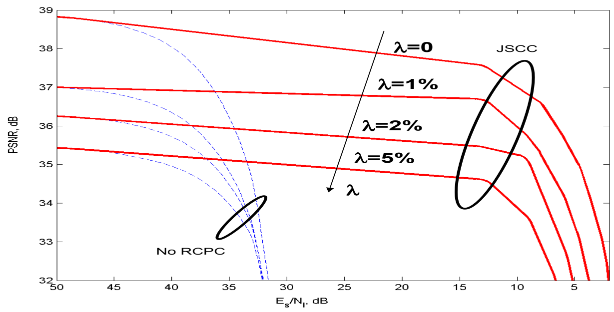

Fig. 3. Performance of H.263+ coded video delivery over heterogeneous wired-to-wireless IP networks using JSCC without an Edge Proxy.

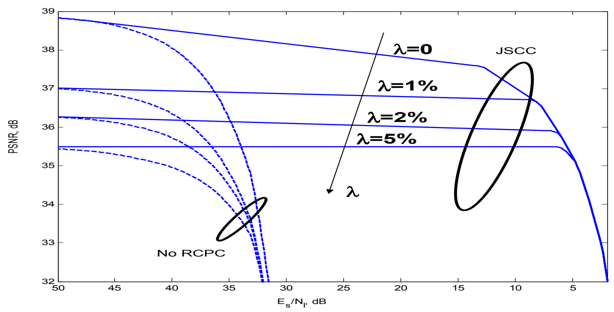

Fig. 4. Performance of H.263+ coded video delivery over heterogeneous wired-to-wireless IP networks using JSCC with an Edge Proxy.

for a set of packet loss rates over the wired IP network with the RS codes and RCPC codes chosen to achieve the overall bitrate budget $R_{tot} = R_s/(R_c^{RCPC} \cdot R_c^{RS}) = 128$ Kbps [2]. In Fig. 3, for a given packet loss rate λ in the wired network, the optimal performance obtainable is demonstrated under the constraint of a fixed wireless transmission rate. It is clear that the RS coding rate has to be adaptively selected with the variation in the corresponding packet loss rate; meanwhile, the RCPC coding has to adapt to the change in the wireless link conditions, E_S/N_I in this case. Clearly, as shown in dashed-line in Fig. 3, for the system employing only adaptive RS codes according to the packet loss rate on the wired network but no RCPC codes on the wireless network, video quality is substantially degraded with increasing bit errors as E_S/N_I decreases. In contrast, the JSCC approach with concatenated RS and RCPC coding provides an effective means to maintain the video quality as network-induced packet-loss and/or bit-error rate increase.

Next, we consider the system with the use of an edge proxy between the wired and wireless IP networks, such that error-control transcoding can be done between the two heterogenous networks requiring different error-control schemes as described in a previous section. With the use of an edge proxy, the corresponding optimal performance obtainable is demonstrated in Fig. 4 under the constraint of the same fixed wireless transmission rate. For comparison, we also present in Fig. 5 the results for the systems with or without the use of an edge proxy under the same transmission rate limit. It clearly demonstrates the substantial improvement using an edge proxy. For example, in the case that packet loss rate over the wired IP network is $\lambda = 5\%$, there is a gain of over 6 dB in wireless channel E_S/N_I for a specified video quality of $PSNR = 34$ dB. This improvement is due primarily to the increase of effective video data throughput due to the error-control transcoding in the

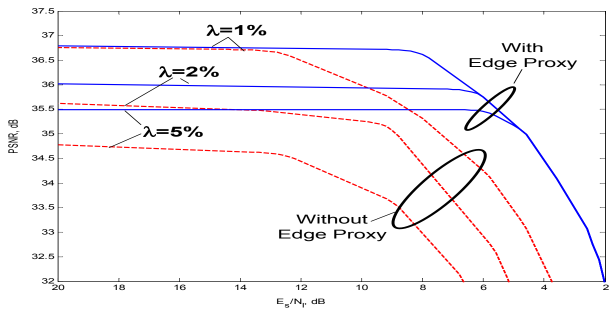

Fig. 5. Performance Improvement with the use of Edge Proxy.

edge proxy. As a result, to meet the same error-protection requirement for both wired and wireless network conditions, a larger effective video data throughput, $R_s = \min\{R_{tot} \cdot R_c^{RS}, R_{tot} \cdot R_c^{RCPC}\}$, is obtained through the use of an edge proxy, compared to $R_s = R_{tot} \cdot R_c^{RS} \cdot R_c^{RCPC}$ for the case without an edge proxy.

6. SUMMARY AND CONCLUSIONS

We have described an end-to-end solution employing an edge proxy operating between the wired and wireless network for packetized H.263+ video over heterogeneous wired-to-wireless IP networks. A JSCC approach employing RS block codes and RCPC codes is studied for the proposed architecture. The results quantitatively demonstrate the requirement for a joint design approach to address the special needs of error-recovery for packet video over the wireless and wired network for acceptable end-to-end quality while exhibiting a graceful pattern of quality degradation in face of dynamically changing network conditions. Furthermore, the results clearly demonstrate the advantage of using an edge proxy in a heterogeneous wired-to-wireless IP networks for improved video quality.

7. REFERENCES

[1] Y. Pei and J. Modestino, "Robust Packet Video Transmission Over Heterogeneous Wired-to-Wireless IP Networks Using ALF Together with Edge Proxies," in *Proc. of EW2002*, Feb. 2002.

[2] Y. Pei and J. Modestino, "Use of Concatenated FEC Coding for Real-Time Packet Video Over Heterogeneous Wired-to-Wireless IP Networks," Submitted to ISCAS2003, May 2003.

[3] C. Bormann, L. Cline, G. Deisher, T. Gardos, C. Maciocco, D. Newell, J. Ott, G. Sullivan, S. Wenger , C. Zhu,, "RTP payload format for the 1998 version of ITU-T Rec. H.263 video (H.263+)." RFC 2429, Oct 1998.

[4] Y. Pei and J. Modestino, "A Joint Source-Channel Coding Approach for Packet Video Transport over Wireless IP Networks," in *Proc. of PV2001*, pp. 41–50, Apr. 2001.

[5] D. Wu, Y. T. Hou, W. Zhu, Y. Q. Zhang and J. M. Peha, "Streaming Video Over the Internet: Approaches and Directions," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 11, pp. 282–300, Mar. 2001.

[6] J. Hagenauer, "Rate-Compatible Punctured Convolutional Codes (RCPC Codes) and their Applications," *IEEE Trans. Commun.*, vol. COM-36, pp. 389–400, April 1988.

[7] Intel Corp., "Video Codec Test Model, TMN8." ftp://standard.pictel.com/video-site/h263plus/ draft13.doc, June 1997.