SYSTEMATIC LOSSY FORWARD ERROR PROTECTION FOR VIDEO WAVEFORMS

Anne Aaron, Shantanu Rane, David Rebollo-Monedero, and Bernd Girod

Information Systems Laboratory, Department of Electrical Engineering
Stanford University, Stanford, CA 94305
{amaaron,srane,drebollo,bgirod}@stanford.edu

ABSTRACT

A novel scheme for error-resilient digital video broadcasting, using Wyner-Ziv coding, is presented in this paper. We apply the general framework of systematic lossy source-channel coding to generate a supplementary bitstream that can correct transmission errors in the decoded video waveform up to a certain residual distortion. The systematic portion consists of a conventional MPEG-coded bitstream, which is transmitted over the error-prone channel without any forward error correction. The supplementary bitstream is a low rate representation of the transmitted video sequence generated using Wyner-Ziv encoding. We use the conventionally decoded error-concealed MPEG video sequence as side information to decode the Wyner-Ziv bits. The decoder combines the error-prone side information and the Wyner-Ziv description to yield an improved decoded video signal. We show how this scheme can be used to build a system which achieves graceful degradation without the need of a layered representation of the video waveform.

1. INTRODUCTION

In typical digital video broadcasting schemes, an MPEG-coded video bitstream is transmitted to multiple receivers over a wireless channel. In order to correct the transmission errors, the source bitstream is generally protected using some form of forward error correction (FEC). The forward error correction scheme, along with decoder-based error concealment ensures the availability of "broadcast quality" video. In this paper, we investigate a novel method for error-resilient video broadcasting which uses Wyner-Ziv coding, instead of conventional forward error correction.

Specifically, we apply the "systematic coding" framework to error resilient MPEG video broadcast. An MPEG-coded video bitstream is transmitted with little or no protection over an error-prone channel. In addition, we transmit a supplementary bitstream generated using Wyner-Ziv coding of the transmitted sequence. We show that a decoder having access to a Wyner-Ziv coded version of the transmitted video sequence, in addition to the error-concealed conventionally decoded sequence as side information, can provide an output with superior visual quality and average PSNR, in the presence of channel errors.

Wyner-Ziv coding refers to lossy compression with side information at the decoder. Achievable rates for this setting were derived in the mid-1970s by Wyner and Ziv [1, 2, 3]. It was proved that the minimum source encoding rate for a given distortion, when the side information is only known to the decoder, is greater than or equal to the rate obtainable when the side information is also available at the encoder. Our implementation of the Wyner-Ziv codec consists of an inner turbo codec and an outer quantization-reconstruction pair.

The Wyner-Ziv problem is closely related to the problem of systematic lossy source-channel coding [4]. In this configuration, an analog source X is transmitted over an analog Channel A without coding. A second encoded version of X is sent over a digital Channel D as enhancement information. The noisy version Y of the original serves as side information to decode the output of Channel D and produce the enhanced version Y^*. The term "systematic coding" has been introduced in extension of systematic error-correcting channel codes to refer to a partially uncoded transmission. Shamai, Verdu, and Zamir establish information theoretic bounds and conditions for optimality of such a configuration in [4].

In conventional forward error correction systems, when the bit error rate of the channel exceeds the correction capability of FEC codes, a "cliff" effect is observed, which results in highly unacceptable video quality. To prevent this, the idea of Priority Encoding Transmission (PET) [5] can be applied, which assigns varying degrees of FEC to different parts of the video bitstream depending upon their relative importance. This approach, which ensures graceful degradation of the image quality in the presence of channel errors, has been exploited by layered video coding schemes [6, 7, 8]. However, these schemes are not used in practice because of the significant rate-distortion penalty associated with a layered video representation. We will describe a scheme using Wyner-Ziv coding that can achieve graceful degradation of the decoded video quality *without the need for a layered representation.*

The remainder of this paper is organized as follows. In Section 2, we explain the lossy forward error protection scheme and describe its fundamental building block, i.e., the Wyner-Ziv codec. In Section 3, we present experimental results of applying this scheme to error-resilient broadcast.

2. SYSTEMATIC LOSSY FORWARD ERROR PROTECTION SCHEME

2.1. Wyner-Ziv Coding for Error-resilient Video Broadcast

The idea is illustrated in Fig. 1, using MPEG video compression as an example. At the transmitter, the input video signal S is compressed independently by an MPEG video coder and a Wyner-Ziv coder. Since the MPEG video bitstream is generated without consideration of the error resilience provided by the Wyner-Ziv coder, we refer to the overall scheme as systematic source-channel coding. The video signal compressed by MPEG and transmitted over an error-prone channel constitutes the systematic portion of the transmission which is augmented by the Wyner-Ziv bitstream. At the receiver, the MPEG bitstream is decoded and transmission errors are concealed, resulting in the decoded video S'. Despite concealment, S' contains some portions of the signal that are degraded by unacceptably large errors. These errors are corrected, up to a residual distortion, by the Wyner-Ziv decoder. The Wyner-Ziv code can be thought of as a second, independent description of the input video S with coarser quantization. To prevent mismatch between the MPEG encoder and decoder, it might be advantageous to use a locally decoded version of the MPEG-compressed video as input to the Wyner-Ziv video encoder, rather than the original video S. Without transmission errors, the Wyner-Ziv description is fully redundant, i.e., it can be regenerated bit-by-bit at the decoder, using the decoded video S'. With transmission errors, Wyner-Ziv bits must be sent to allow an error-free reconstruction of the coarser second description, employing the decoded video signal S' as side information. The error correction capabilities of the Wyner-Ziv bitstream can be simultaneously used to protect the Wyner-Ziv bits against transmission errors. The coarser second description and side information S' are optimally combined to yield an improved decoded video signal S^*. In portions where the waveform S' is not affected by transmission errors, S^* will be essentially identical to S'. However, in portions of the waveform where S' is substantially degraded by transmission errors, the second coarser representation transmitted at very low bit-rate in the Wyner-Ziv bitstream limits the maximum degradation that can occur. Instead of the error-concealed decoded signal S', the signal S^* at the output of the Wyner-Ziv encoder is fed back to the MPEG decoder to serve as a more accurate reference frame for decoding of further frames.

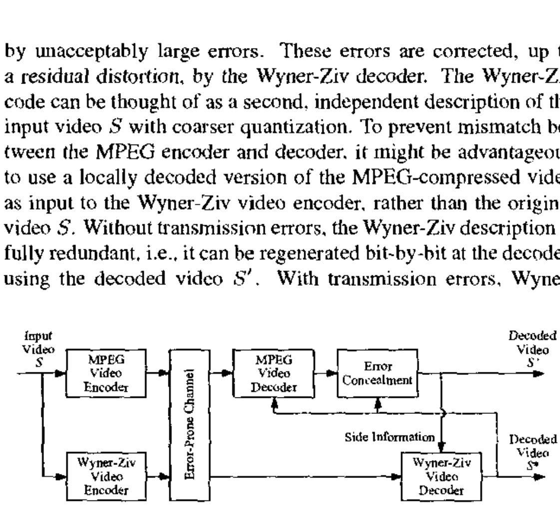

Fig. 1. Wyner-Ziv bitstream uses decoded error-concealed video waveform as side information in a systematic lossy source-channel set-up.

2.2. Wyner-Ziv Codec

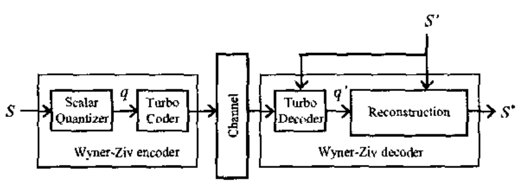

Fig. 2. Wyner-Ziv codec consists of a inner turbo codec and an outer quantization-reconstruction pair.

The Wyner-Ziv encoder and decoder can be seen in Fig. 2. A similar Wyner-Ziv codec structure was presented in [9, 10], for an asymmetric video compression system employing intraframe encoding but interframe decoding. Note, however, that the present codec is designed for a broadcast application and hence does not employ feedback, unlike these earlier schemes. At the encoder, each pixel of the video sequence S is quantized using a dithered scalar quantizer of Q levels. The quantized symbols, q, corresponding to a frame are grouped together to form the input block to the turbo encoder. The turbo encoder, composed of two constituent systematic convolutional encoders, generates parity bits which are sent to the Wyner-Ziv decoder. Note that if there are no transmission errors, these parity bits do not give any additional information. This is similar to the inefficiency of traditional forward error correction when there are no errors.

The Wyner-Ziv decoder uses the parity bits and the side information S' (the decoded sequence with concealed transmission errors) to decode the quantized symbols q'. The reconstruction block takes a decoded symbol q_i' and the corresponding side information to form a better reconstruction S_i^*. If the side information is within the interval of the decoded symbol, S_i^* takes the value of the side information. If it is outside the interval, S_i^* is calculated by taking the conditional expectation (assuming a Laplacian residual model between the encoded signal and the side information) given the side information and the decoded interval. Assuming that the decoded symbol is correct, the Wyner-Ziv codec limits the distortion of the output sequence up to a maximum distortion determined by the quantizer coarseness Q.

2.3. Embedded Wyner-Ziv Codec

The trade-off between distortion due to transmission errors and Wyner-Ziv bit-rate can be exploited to construct an embedded Wyner-Ziv codec that achieves graceful degradation of the decoded video when the error rate of the channel increases. An example of such a system is shown in Fig. 3 for the case of 2 quality levels. Wyner-Ziv encoder A employs a coarser representation that is embedded in the finer representation of Wyner-Ziv encoder B. Since Wyner-Ziv encoder A has a coarser quantizer, the bitstream is easier to decode and, therefore, has stronger error protection capabilities. It is decoded first, using decoded video S' as side information to yield improved decoded video S^*. If the transmission errors are not too severe, then the Wyner-Ziv stream B can also be decoded, using the decoded probabilities of the symbols from Wyner-Ziv decoder A, P_A, and the side information S'. This yields a further improved decoded video signal S^{**}. Thus, graceful degradation of the decoded video quality is achieved without requiring a layered coding scheme.

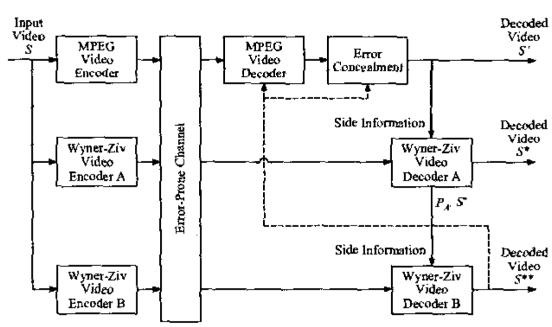

Fig. 3. Graceful degradation of decoded video quality can be achieved using an embedded Wyner-Ziv codec.

3. EXPERIMENTAL RESULTS

3.1. Wyner-Ziv Coding for Forward Error Protection

We implemented the scheme proposed in Section 2.1 for standard MPEG-compressed CIF sequences. For the simulations we assume that the channel errors lead to lost macroblocks. We simulate a random macroblock loss rate of 1%. The decoder performs previous frame error-concealment on lost macroblocks to generate the sequence S', which serves as side information for the Wyner-Ziv decoder. The experimental set-up assumes error-free transmission of the Wyner-Ziv bits.

The CIF video format produces for every frame a symbol block of length $288 \times 352 = 101376$ which is the input to the turbo encoder. To vary the level of error protection, we change the number of quantization levels ($Q \in \{4, 16\}$) and control the amount of parity bits sent to the decoder. Note that for this application there is no feedback in the system; so for a given simulation set-up the Wyner-Ziv transmission rate is fixed.

As explained in Section 2.1, the Wyner-Ziv decoder optimally combines the side information S', and the Wyner-Ziv bits to generate a signal S^* of improved visual quality. Fig. 4 shows the the PSNR vs. frame number of the *Carphone* sequence MPEG-encoded at 1 Mbps and protected by increasing amounts of Wyner-Ziv bits. It can be observed that our scheme limits the error propagation caused by imperfect concealment of channel errors, and achieves acceptable PSNR levels. For low rates of the Wyner-Ziv bitstream, the error protection capability breaks down when the errors are too severe. As expected, the average PSNR is largest when the rate of the Wyner-Ziv bitstream is largest (1 bpp in our simulations). We note that this rate is high at present because our pixel domain Wyner-Ziv coder does not exploit any spatial correlation in the input video sequence, thus, the bit rate is not yet comparable to traditional FEC. Upon close inspection, a few isolated pixel errors can be observed (approx. 100 pixels per CIF frame) in the reconstructed video frame. These are due to the presence of residual errors in the turbo decoder. For high resolution video, which is typical for video broadcast, isolated pixel errors are not very evident to the viewer.

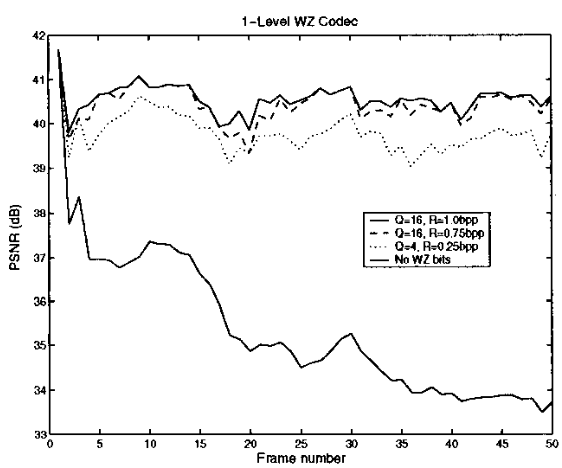

Fig. 4. PSNR of the decoded frames improve as the amount of Wyner-Ziv bits increase.

3.2. Wyner-Ziv Coding Applied to Temporally Subsampled Source Sequence

In this section we discuss results obtained when the supplementary bitstream consists of a Wyner-Ziv representation of a temporally subsampled version of the source sequence, instead of the entire sequence itself. This reduces the effective Wyner-Ziv bit-rate (from 2 bpp with no temporal subsampling, to $2/6 = 0.33$ bpp with Wyner-Ziv coding applied once every 6 frames) at the expense of an increase in the average distortion due to transmission errors. Fig. 5 shows how the average PSNR degrades with increase in the temporal subsampling factor applied to the source sequence before Wyner-Ziv coding. Note that the degradation is not very severe compared to that observed when Wyner-Ziv decoding is not used. The tradeoff between temporal subsampling versus Wyner-Ziv bit-rate should therefore be an important consideration in the optimization of the forward error protection system.

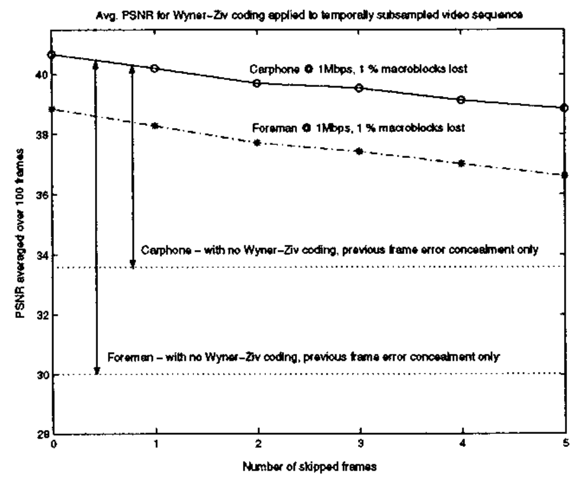

Fig. 5. Trade-off between Wyner-Ziv bit-rate and average PSNR, when Wyner-Ziv coding is applied to a temporally subsampled version of the source sequence.

3.3. Embedded Wyner-Ziv Coding

We simulated a 2-level embedded Wyner-Ziv coder. The coarser representation of Wyner-Ziv codec A is formed using a 4-level uniform quantizer which is embedded in a 16-level uniform quantizer used for Wyner-Ziv codec B. Once codec A decodes the symbols correctly, it passes the decoded probabilities of the symbols to codec B.

A plot of the PSNR vs. frame number for the *Carphone* sequence can be seen in Fig. 6. For this plot, Wyner-Ziv stream A has a fixed bit-rate of $R_4 = 0.25$ bpp, which is enough to correctly decode the coarse representation of all the frames. The bit-rate of stream B, R_{16}, is varied across the plots. As it can be seen, $R_{16} = 0.375$ bpp can successfully decode the finer representation of all the frames. At $R_{16} = 0.25$ bpp, codec B fails decoding for some frames and the PSNR falls to the level of that of Wyner-Ziv codec A.

In Fig. 7, we compare the 2-level embedded codec with the 1-level codec with the same bit-rate. For the 1-level codec with 16

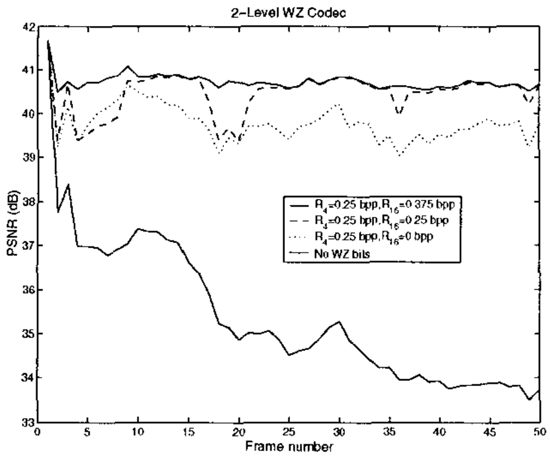

Fig. 6. A 2-level embedded Wyner-Ziv coding of the transmitted sequence ensures graceful degradation of the decoded video quality without the need for a layered representation of the source sequence.

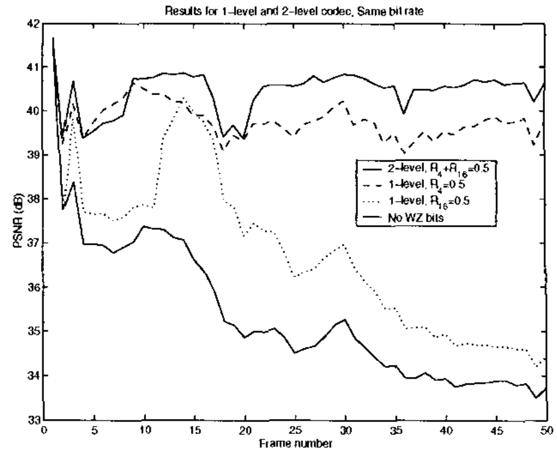

Fig. 7. Comparison of the performance of forward error protection using basic Wyner-Ziv coding of the source sequence, and 2-level (embedded) Wyner-Ziv coding

quantization levels, the given bit-rate is not enough to correctly decode most of the frames, so the PSNR falls. For the 1-level codec with 4 quantization levels, the given bit-rate is more than enough to correctly decode all the frames but the PSNR improvement is limited by the coarse quantization. However, for the 2-level codec, the bit-rate is used efficiently. When the side information is good, it is able to decode the finer representation. When the side information is bad, it is still able to decode the coarse representation, thus, avoiding a drop in PSNR.

4. CONCLUSIONS

In this paper we have explored the application of Wyner-Ziv coding to error resilient digital video broadcasting. Experimental results show that a supplementary bitstream generated using Wyner-Ziv coding of the source sequence can be used to correct transmission errors in the transmitted video signal, up to a certain residual distortion. Since it allows for some distortion in the case of channel errors, the above scheme can potentially achieve a much lower bit-rate than a conventional channel coder which protects the bits produced by the source coder. Equivalently, we can achieve stronger error protection at the same bit-rate, if we allow for higher distortion. Further, by using an embedded Wyner-Ziv codec, graceful degradation of the decoded video quality can be achieved without the requirement of a layered representation of the video source.

Refinements in the decoder reconstruction function, a transform domain Wyner-Ziv codec, and improved quantization schemes based on the conditional entropy of the source given the side information, constitute the focus of our future work.

5. REFERENCES

[1] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. IT-22, no. 1, pp. 1–10, Jan. 1976.

[2] A. Wyner, "On source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. IT-21, no. 3, pp. 294–300, May 1975.

[3] A. Wyner, "The rate-distortion function for source coding with side information at the decoder-II: General sources," *Information and Control*, vol. 38, pp. 60–80, 1978.

[4] S. Shamai, S. Verdu, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 564–579, Mar. 1998.

[5] A. Albanese, J. Blomer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoding transmission," *IEEE Transactions on Information Theory*, vol. 42, no. 6, pp. 1737–1744, Nov. 1996.

[6] M. Gallant and F. Kossentini, "Rate-distortion optimized layered coding with unequal error protection for robust internet video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 3, pp. 357–372, Mar. 2001.

[7] A. Mohr, E. Riskin, and R. Ladner, "Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 819–828, June 2000.

[8] U. Horn, K. Stuhlmüller, M. Link, and B. Girod, "Robust internet video transmission based on scalable coding and unequal error protection," *Image Communication, Special Issue on Real-time Video over the Internet*, vol. 15, no. 1-2, pp. 77–94, Sept. 1999.

[9] A. Aaron, R. Zhang, and B. Girod, "Wyner-Ziv coding of motion video," in *Proc. Asilomar Conference on Signals and Systems*, Pacific Grove, California, Nov. 2002.

[10] A. Aaron, S. Rane, R. Zhang, and B. Girod, "Wyner-Ziv coding for video - Applications to compression and error resilience," in *Proc. IEEE Data Compression Conference*, Snowbird, Utah, Mar. 2003, pp. 93–102.