SYSTEMATIC LOSSY FORWARD ERROR PROTECTION FOR ERROR-RESILIENT DIGITAL VIDEO BROADCASTING - A WYNER-ZIV CODING APPROACH

Shantanu Rane, Anne Aaron and Bernd Girod

Information Systems Laboratory, Stanford University, Stanford, CA 94305
{srane,amaaron,bgirod}@stanford.edu

ABSTRACT

We present a practical scheme for error-resilient digital video broadcasting, using the Wyner-Ziv coding paradigm. 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-2 bitstream, which is transmitted over the error-prone channel without forward error correction. The supplementary bitstream is a low rate representation of the transmitted video sequence generated using Wyner-Ziv encoding. We use the received error-prone MPEG-2 prediction error signal 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 describe a system that uses an embedded Wyner-Ziv codec to achieve graceful quality degradation without the need for a layered representation of the video waveform.

1. INTRODUCTION

In typical digital video broadcasting systems, an MPEG-coded video bitstream is transmitted to multiple receivers. To correct 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. However, when the channel error rate exceeds the error correction capability of the FEC codes, the video quality degrades rapidly, leading to the undesirable "cliff" effect. In this paper, we investigate a novel method for error-resilient video broadcasting which uses Wyner-Ziv coding, instead of conventional forward error correction.

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 encoding rate for a source sequence X, for a given distortion, when the side information Y 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. Zamir [4] showed that the rate loss associated with ignoring the side information at the encoder is upper-bounded by the minimax capacity of an additive noise channel, where the side information Y is thought to be a noisy version of X. Even with this rate loss, the encoding rate with Wyner-Ziv coding is lower than that achieveable in the conventional non-distributed case owing to the correlation between X and Y. In general, the Wyner-Ziv codec consists of an inner channel codec and an outer quantization-reconstruction pair. For this paper, these functions are performed by a Reed-Solomon codec and an MPEG-2 quantizer respectively.

The Wyner-Ziv problem is closely related to the problem of systematic lossy source-channel coding [5]. 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 as an extension of systematic error-correcting channel codes to refer to a partially uncoded transmission. Shamai, Verdú, and Zamir established information theoretic bounds and conditions for optimality of such a configuration in [5]. The systematic coding framework was used by Pradhan and Ramchandran [6] for enhancing the quality of images corrupted by additive white Gaussian noise, using digital side information

The undesirable FEC "cliff" effect can be prevented by using Priority Encoding Transmission (PET) [7] 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 [8, 9, 10]. However, these schemes are not used in practice because of the inefficient rate-distortion performance of layered video coding. In this paper, we describe a scheme using Wyner-Ziv coding that can achieve graceful degradation of the decoded video quality *without the need for a layered representation*.

First results of applying simple pixel-domain Wyner-Ziv coding for error resilient video broadcasting were presented in our own work [11, 12]. However, since the Wyner-Ziv codec did not exploit any spatial or temporal correlation in the video sequence, these schemes could not compete with the low bitrates achieved by traditional broadcast systems which use FEC. We presented substantially improved results using a hybrid video codec in conjunction with Reed-Solomon codes [13]. This scheme required an extra video encoder inside the Wyner-Ziv decoder and thus, had high computational complexity. In this paper, we describe a practically implementable system which is fully backward-compatible with legacy broadcast systems and operates with negligible complexity overhead.

Recently, other distributed video coding schemes have been proposed, albeit for different applications. In particular, Puri and Ramchandran presented the PRISM architecture [14], using uses channel coding concepts to achieve flexible distribution of computational complexity between the encoder and decoder along with robustness to predictive mismatch. Aaron and Girod presented a Wyner-Ziv video codec ideally suited for applications like cameraphones which require low complexity encoding, where individual frames are en-

This work is supported in part by NSF Grant No. CCR-0310376 and C.V.Starr Southeast Asian Fellowship

coded independently but decoded conditionally [15, 16]. Sehgal *et al.* [17, 18] used coset codes to design a video codec immune to predictive mismatch.

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 broadcasting.

2. SYSTEMATIC LOSSY FORWARD ERROR PROTECTION SCHEME

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

The concept of systematic lossy forward error protection 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'. Even after concealment, S' contains some portions that are degraded by unacceptably large errors. These errors are corrected, up to a certain 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, but with coarser quantization. To prevent mismatch between the MPEG encoder and decoder, it is 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 then fully redundant, i.e., it can be regenerated bit-by-bit at the decoder, using the decoded video S'.

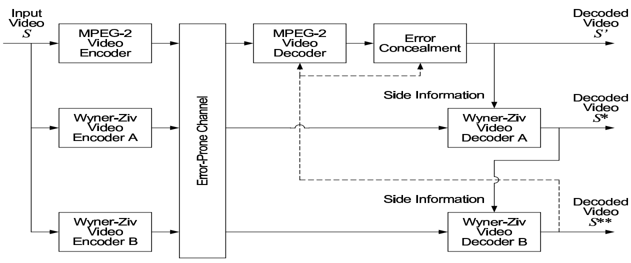


Fig. 1. Wyner-Ziv decoder uses decoded error-concealed video waveform as side information in systematic lossy source-channel setup.

With transmission errors, Wyner-Ziv bits must be sent to allow 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 combined to yield an improved decoded video signal S^*. In portions where the waveform S' is not affected by transmission errors, S^* is 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 bitrate 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. The systematic scheme described in Fig. 1 is compatible with systems already deployed, such as MPEG-2 digital TV broadcasting systems. The Wyner-Ziv bitstreams can be ignored by legacy systems, but would be exploited by new receivers.

2.2. Wyner-Ziv Codec

We now describe a practical Wyner-Ziv codec constructed from well-understood components viz., quantizers, entropy coders, and a Reed-Solomon (RS) codec. As shown in Fig. 2, The Wyner-Ziv encoding process consists of two stages:

1. A coarse quantizer operates on the quantized transformed prediction error signal generated by the conventional MPEG-2 encoder. The Wyner-Ziv description now consists of the entropy coded output of the coarse quantizer along with the motion vectors and mode decisions inherited from the conventional MPEG-2 encoder.

2. The resulting bitstream is input to a channel coder which applies systematic RS codes with byte-long symbols, across the slices of an entire frame. *Only the RS parity symbols* are then transmitted to the receiver, and these consitute the Wyner-Ziv bitstream. *The systematic portion of the RS encoder output is discarded.*.

If there are no transmission errors, the RS parity symbols do not provide any additional information. This is similar to the inefficiency of traditional forward error correction when there are no errors. When transmission errors occur, the decoder generates a coarsely quantized version of the received prediction error signal. This coarse version is actually an error-prone copy of the Wyner-Ziv description, which serves as side information for the RS decoder. The RS decoder uses the parity symbols and error-prone Wyner-Ziv description to obtain the error-free Wyner-Ziv description. Since the location of the lost slices is known, the RS decoder can perform *erasure decoding* across the error-prone slices. A fallback mechanism substitutes the lost slices in the main video sequence with their correct but coarser versions. After decoding, the coarse fallback causes some prediction mismatch which propagates to the subsequent frames, but visual examination of the decoded sequence shows that this small error is imperceptible. Thus, the receiver obtains a video sequence of superior visual quality. This system includes FEC as a special case, if the "coarse" quantizer uses the same quantization parameter as the main MPEG-2 encoder. Note that this scheme applies Wyner-Ziv decoding to the *received* prediction error signal, as opposed to generating side information by re-encoding the output of the MPEG decoder, as was done in [13]. Therefore, the decoder implementation is very simple, with the coarse quantizer Q_1 and extra entropy coding adding a negligible complexity overhead with respect to conventional FEC techniques.

2.3. Embedded Wyner-Ziv Codec

The trade-off between the distortion due to transmission errors and Wyner-Ziv bitrate can be exploited to construct an embedded Wyner-Ziv code 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. 1 for the case of 2 quality levels.

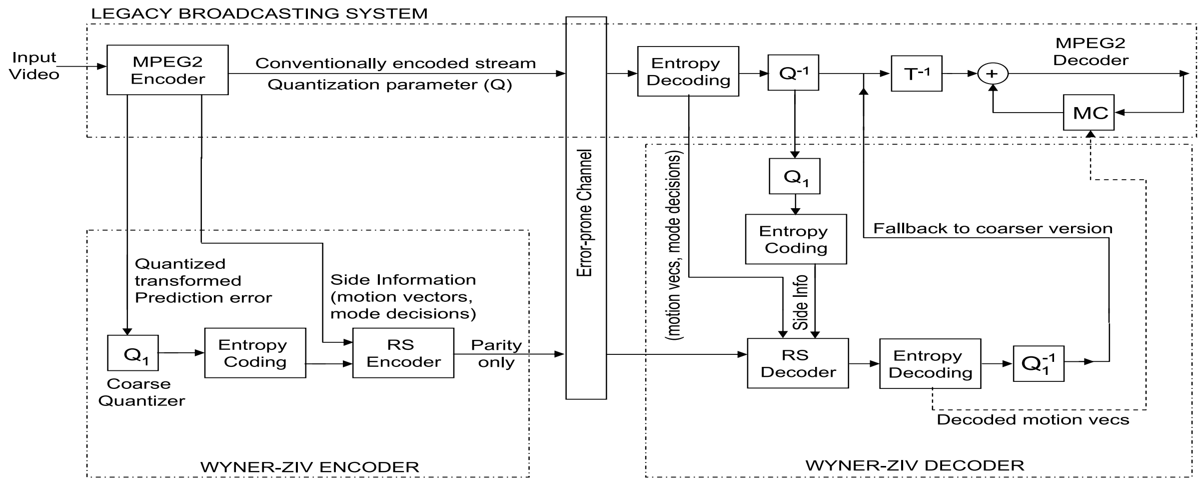


Fig. 2. Implementation of systematic forward error protection by combining MPEG coding and Reed-Solomon codes across slices.

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, its 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, possibly using the side information S'. This yields a further improved decoded video signal S^{**}. Thus, graceful degradation of decoded video quality is achieved without a layered coding scheme.

3. EXPERIMENTAL RESULTS

3.1. Wyner-Ziv Coding for Forward Error Protection

The performance of our system over a range of symbol error rates is shown for the *Foreman CIF* sequence in Fig. 3, along with traditional FEC for comparison. Each plot corresponds to a main MPEG-2 bitstream, with the *I-B-B-P* frame structure, encoded at 2 Mbps, with an additional 222 Kbps worth of RS parity information. As it can be seen, the PSNR for the system using conventional FEC drops rapidly beyond a symbol error probability of 3×10^{-5}. The forward error protection (FEP) scheme also uses 222 Kbps worth of RS parity symbols, but the RS code is applied to coarser versions of the original video sequence, encoded at 1 Mbps and 500 Kbps respectively. Therefore, FEP can use stronger RS protection at the same bitrate, and can provide acceptable video quality over a wider range of symbol error rates. For low symbol error probability, there is a small reduction in the PSNR owing to the coarser quantization employed in FEP. But at higher symbol error probabilities, the FEP schemes significantly outperform traditional FEC.

3.2. Embedded Wyner-Ziv Coding

In Figs. 3 and 4, we compare the 2-level embedded codec with the 1-level codec at the same bitrate. In the 2-level codec, Wyner-Ziv stream A consists of 166 Kbps of parity information corresponding to a coarsely quantized version of the video sequence, encoded at 500 Kbps. The Wyner-Ziv stream B consists of 56 Kbps of parity information, corresponding to a finely quantized version encoded at 1 Mbps. We compare this with a 1-level codec in which the Wyner-Ziv stream contain 222 Kbps of parity information applied to (1) the coarsely quantized description alone, and (2) the finely quantized description alone. For the 1-level codec with finer quantization, the Wyner-Ziv bitrate is too low to correctly decode most of the frames, so the PSNR falls rapidly. For the 1-level codec with coarser quantization, the Wyner-Ziv bitrate is sufficient to correctly decode all the frames but the PSNR is affected by the coarse quantization. However, for the 2-level codec, the bitrate is used efficiently. When a video frame has few error-prone slices, the finer representation can be decoded. When a large number of slices are in error, it is still possible to decode the coarser representation, thus avoiding a large drop in PSNR. Since quality degradations due to transmission errors are localized within the sequence, PSNR values averaged over the entire sequence are of limited value. A better appreciation of the clear advantages of embedded Wyner-Ziv coding can be obtained from the traces in Fig. 4 and by comparing the visual quality of the decoded video, as shown in Fig. 5.

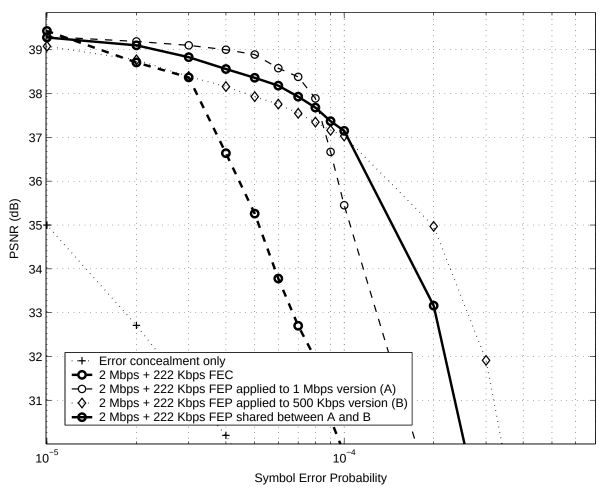


Fig. 3. Forward error protection (FEP) achieves graceful degradation of the decoded video quality, compared to traditional FEC. Embedding two FEP streams exploits the tradeoff between error-resilience and the Wyner-Ziv bitrate, to obtain graceful quality degradation without a layered video codec.

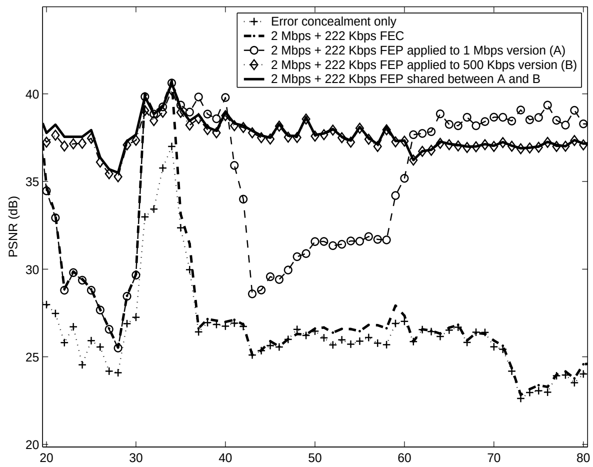
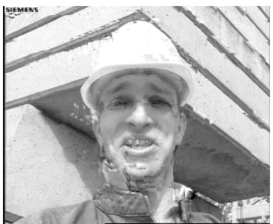
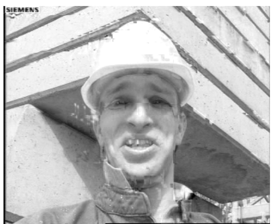


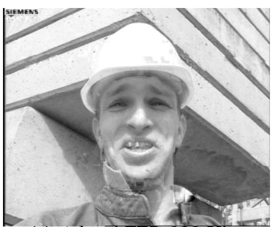
Fig. 4. Depending upon the quality of the side information, the PSNR of the embedded Wyner-Ziv coding scheme remains greater than or equal to that of the single layer scheme with coarsely quantized Wyner-Ziv description. All traces are at symbol error probability of 10^{-4}.



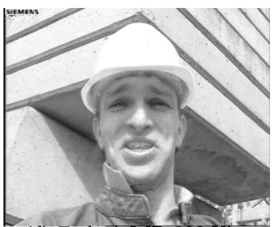
(a) Error Concealment Only



(b) FEC, 222 Kbps



(c) 1-level FEP, 222 Kbps



(d) Embed. WZ, 222 Kbps

Fig. 5. Substantial improvement in decoded video quality is observed by using embedded Wyner-Ziv coding. Decoded frame #55 from Fig. 4 is shown here for visual comparison.

4. CONCLUSIONS

In this paper we presented a practical scheme which applies Wyner-Ziv coding ideas for 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 received video signal, up to a certain residual distortion. In return for some impreceptible residual distortion in the case of channel errors, the above scheme can potentially achieve a much lower bitrate than a conventional channel coder which protects the bits produced by the source coder. Equivalently, we can achieve stronger error protection at the same bitrate, if we allow for higher distortion. This trade-off has been used to build a embedded Wyner-Ziv codec, that achieves graceful degradation of the decoded video quality without needing a layered representation of the video source.

5. REFERENCES

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

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

[3] A. Wyner, "The rate-distortion function for source coding with side information at the decoder—II: General sources," *Inform., Contr.*, vol. 38, no. 1, pp. 60–80, July 1978.

[4] R. Zamir, "The rate loss in the Wyner-Ziv problem," *IEEE Trans. Inform. Theory*, vol. 42, no. 6, pp. 2073–2084, Nov. 1996.

[5] S. Shamai, S. Verdú, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 564–579, Mar. 1998.

[6] S. S. Pradhan and K. Ramchandran, "Enhancing analog image transmission systems using digital side information: a new wavelet based image coding paradigm," in *Proc. IEEE Data Compression Conference*, Snowbird, UT, Mar. 2001, pp. 63–72.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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, UT, Mar. 2003, pp. 93–102.

[12] A. Aaron, S. Rane, D. Rebollo-Monedero, and B. Girod, "Systematic lossy forward error protection for video waveforms," in *Proc. IEEE International Conference on Image Processing*, Barcelona, Spain, Sept. 2003.

[13] S. Rane, A. Aaron, and B. Girod, "Systematic lossy forward error protection for error resilient digital video broadcasting," in *Proc. SPIE Visual Communications and Image Processing*, San Jose, CA, Jan. 2004.

[14] R. Puri and K. Ramchandran, "PRISM: A new robust video coding architecture based on distributed compression principles," in *Proc. Allerton Conference on Communication, Control, and Computing*, Allerton, IL, Oct. 2002.

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

[16] A. Aaron, E. Setton, and B. Girod, "Towards practical Wyner-Ziv coding of video," in *Proc. IEEE International Conference on Image Processing*, Barcelona, Spain, Sept. 2003.

[17] A. Sehgal and N. Ahuja, "Robust predictive coding and the Wyner-Ziv problem," in *Proc. IEEE Data Compression Conference*, Snowbird, UT, Mar. 2003, pp. 103–112.

[18] A. Sehgal, A. Jagmohan, and N. Ahuja, "A causal state-free video encoding paradigm," in *Proc. IEEE International Conference on Image Processing*, Barcelona, Spain, Sept. 2003.