Systematic Lossy Forward Error Protection for Error-Resilient Digital Video Broadcasting

Shantanu Rane, Anne Aaron and Bernd Girod

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

ABSTRACT

We present a novel 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-coded 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 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. Our results indicate that, over a large range of channel error probabilities, this scheme yields superior video quality when compared with traditional forward error correction techniques employed in digital video broadcasting.

Keywords: Wyner-Ziv coding, systematic lossy source-channel coding, forward error protection, side information decoding.

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 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 protection.

Specifically, we apply the "systematic coding" framework to error-resilient MPEG video broadcasting. An MPEG-coded video bitstream is transmitted with little or no protection over an error-prone channel. However, 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.

Please send correspondence to Shantanu Rane. Email: srane@stanford.edu, Phone: 650-724-3647

2. RELATED WORK

Wyner-Ziv coding refers to lossy compression with correlated side information at the decoder. Achievable rates for this setting were derived in the mid-1970s by Wyner and Ziv [1–3]. 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 the ignorance of 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 a H.264 video codec 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, Verdu, and Zamir establish 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 analog image transmission systems, using digital side information

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) [7] 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 [8–10]. 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.*

First results of applying simple pixel-domain Wyner-Ziv coding for error resilient video broadcasting were presented in our own work in [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. Recently, other distributed video coding schemes have been proposed, albeit for different applications. In particular, Puri and Ramchandran presented their PRISM architecture [13], which utilizes channel coding concepts to achieve a flexible distribution of computational complexity between the encoder and decoder along with robustness to predictive mismatch. In an independent parallel development, Sehgal et al. [14, 15] used coset codes to design a video codec immune to predictive mismatch, which is only 1 to 2.5 dB worse than the baseline H.26L codec.

The remainder of this paper is organized as follows. In Section 3, we explain the lossy forward error protection scheme and describe its fundamental building block, i.e., the Wyner-Ziv codec. In Section 4, we present experimental results of applying this scheme to error-resilient broadcasting, before the concluding remarks in Section 5.

3. SYSTEMATIC LOSSY FORWARD ERROR PROTECTION

3.1. 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'. Even after concealment, S' contains some portions of the signal 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 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'. With transmission errors, Wyner-Ziv bits must be sent to allow an error-free

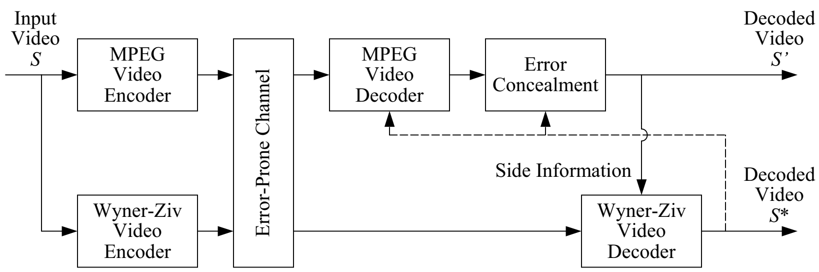

Figure 1. Wyner-Ziv bitstream uses decoded error-concealed video waveform as side information in a systematic lossy source-channel set-up.

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^* 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. The systematic scheme described in Fig. 1 is compatible with systems already deployed, such as MPEG-2 digital television broadcasting systems. The Wyner-Ziv bitstreams can be ignored by legacy systems, but would be exploited by new receivers.

3.2. Wyner-Ziv Codec

The Wyner-Ziv codec in earlier versions of the systematic lossy forward error protection (FEP) schemes [11, 16], used an inner turbo codec and an outer quantization-reconstruction pair. As shown in Fig. 2, the Wyner-Ziv codec in the present scheme uses a hybrid video codec in which the video frames are divided into the same slice structure as that used in the MPEG video coder, but are encoded with coarser quantization. The bitstream from this second encoder is input to a channel coder which applies systematic Reed-Solomon (RS) codes across the slices of an entire frame, as shown in Fig. 3. *only the RS parity symbols* are transmitted to the receiver. If there are no transmission errors, these parity symbols do not give any additional information. This is similar to the inefficiency of traditional forward error correction when there are no errors.

When transmission errors occur at the decoder, a "coarse" video encoder inside the Wyner-Ziv decoder (which is essentially a replica of the coarse encoder inside the Wyner-Ziv encoder), re-encodes the conventionally decoded error-concealed video S'. The output of this coarse encoder is 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

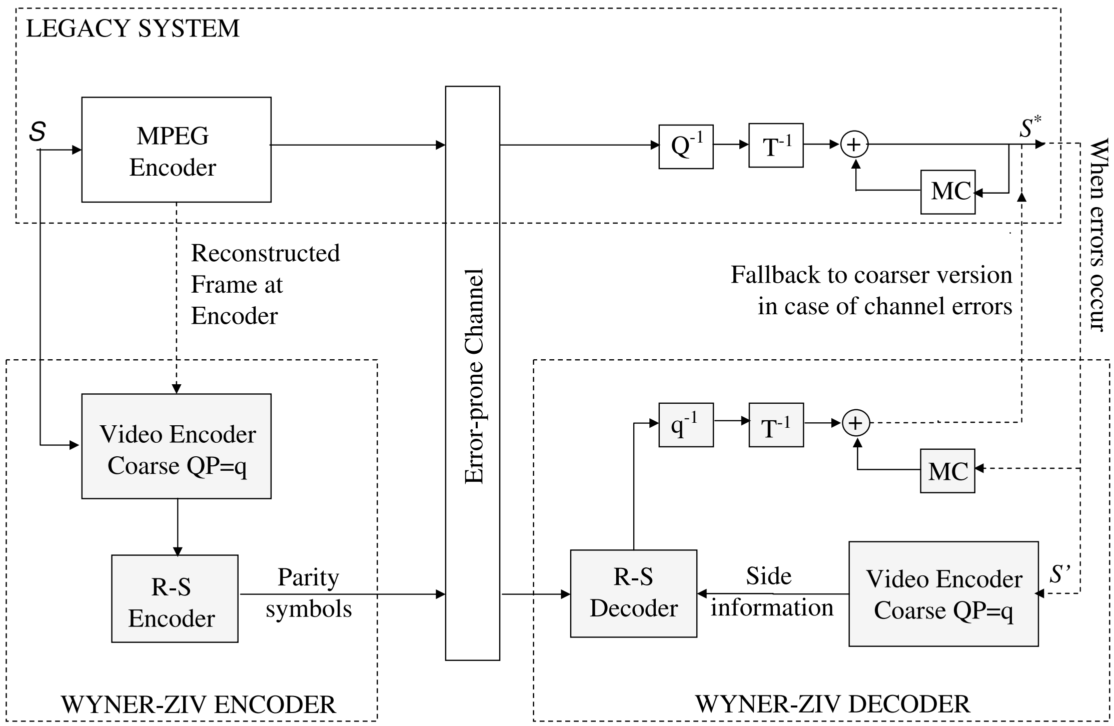

Figure 2. Wyner-Ziv codec in the proposed systematic lossy forward error protection (FEP) scheme

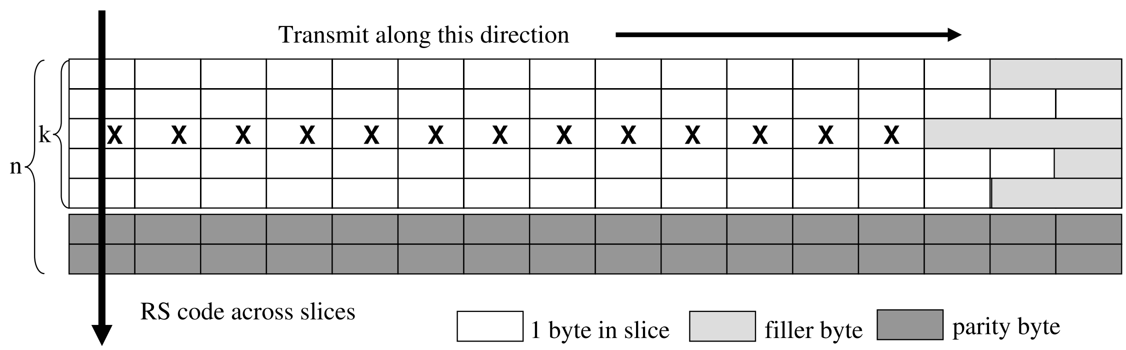

Figure 3. Reed Solomon codes applied across slices. Knowledge of the location of erroneous slices enables erasure decoding.

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. 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 S^* of superior visual quality. This system includes FEC as a special case, which is realized by forcing both encoders in Fig. 2 to use the same quantization parameter.

Note that, to prevent prediction mismatch between the two coarse video encoders housed inside the Wyner-Ziv codec, it is necessary to ensure that they use a common reference frame for predictive coding. This is done by requiring the coarse encoder inside the Wyner-Ziv encoder to use the locally decoded reference frame from the "main" MPEG encoder, as a reference frame for predictive coding.

Further, the fact that the main MPEG encoder and the coarse encoder in the Wyner-Ziv encoder differ only in the quantization parameter, can be used to streamline the encoder implementation. This can be done by using a feature known as "redundant slices" which is already available in the H.264 video coding standard [17, 18]. This feature allows encoding of one or more redundant versions of a slice, with each version using different quantization parameters. Hence, the functions of the main MPEG encoder and one or more Wyner-Ziv encoders can be implemented with a single H.264 video codec.

4. EXPERIMENTAL RESULTS

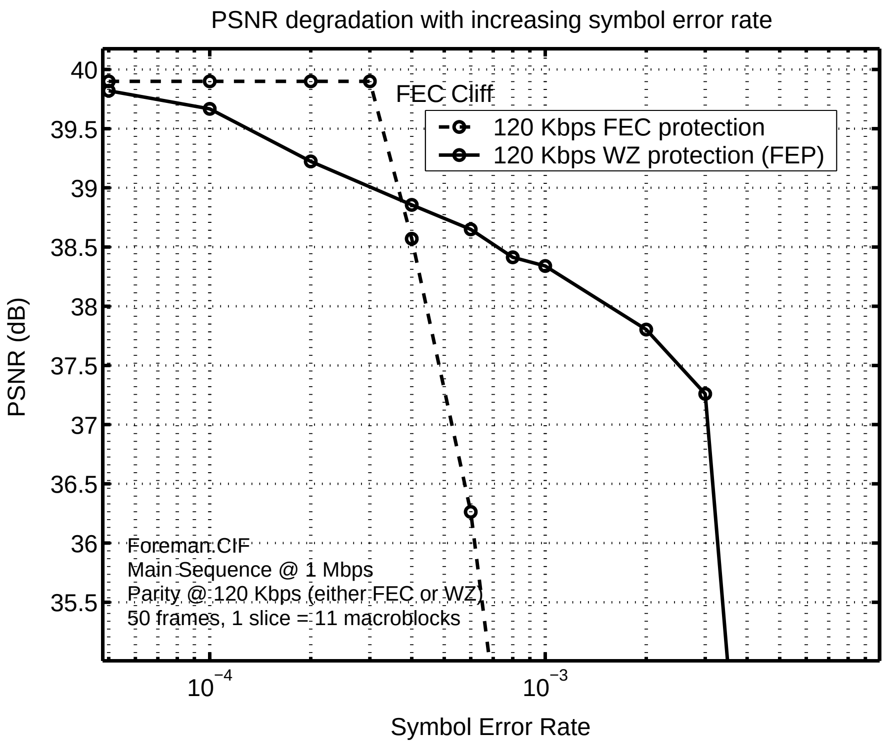

Figure 4. Proposed systematic lossy forward error protection (FEP) scheme outperforms conventional FEC when the symbol error probability increases.

We implemented the scheme proposed in Section 3.1 for standard MPEG-compressed CIF sequences. As in typical broadcast scenarios, the video frames are divided into slices for error-resilience. In our simulations, 11 consecutive macroblocks are placed into one slice, so that there are 36 slices in a CIF frame. Transmission errors corrupt the bitstream and result in incorrectly decoded slices which are concealed using previous frame error concealment. The frame containing the erroneous slices is presented as side-information to the Wyner Ziv decoder.

The Wyner-Ziv encoder contains a second hybrid video encoder which encodes the input video sequence at a fraction of the bit-rate used by the main MPEG video coder, because it uses coarser quantization. To this coarse representatio, RS codes (using byte-long symbols) are applied across the slices of an entire video frame. The parity symbols generated by the RS coder are then transmitted as the supplementary bitstream, which we refer to as the "Wyner-Ziv bitstream".

Fig. 4 shows a typical comparison between the systematic lossy FEP scheme and traditional FEC. The systematic portion of the bitstream is the Foreman sequence encoded at 1 Mbps (with Quantization Parameter QP=11). The Wyner-Ziv description is the Foreman sequence encoded at 270 Kbps (QP=24). For traditional

FEC, a (40,36) RS code is applied to the MPEG bitstream. For FEP, a (52,36) RS code is applied to the Wyner-Ziv description. Note that, given a fixed bit-rate for error protection, the coarser MPEG bitstream can use stronger RS protection than could be used for FEC of the main MPEG bitstream. In fact, in this example, both FEC and FEP use 120 Kbps worth of parity information. As seen in Fig. 4, the error correction capability of FEC breaks down at a symbol error probability of 3×10^{-4}. Beyond this error probability, the decoder resorts to previous frame error concealment, the PSNR drops and the visual quality degrades rapidly (Fig. 5), producing the FEC "cliff". However, FEP continues to provide acceptable video quality up to a higher error probability of 3×10^{-3}. Beyond this error probability, the channel errors overwhelm the Wyner-Ziv decoder and a FEC-like cliff appears. For error probabilities to the left of the FEC cliff, we observe that FEC is marginally better than FEP, owing to the propagation of quantization errors from the coarse fallback mechanism employed by FEP. Visual examination has shown these to be imperceptible. For error probabilities to the right of the FEC cliff, FEP significantly outperforms FEC *while using the same bit-rate*.

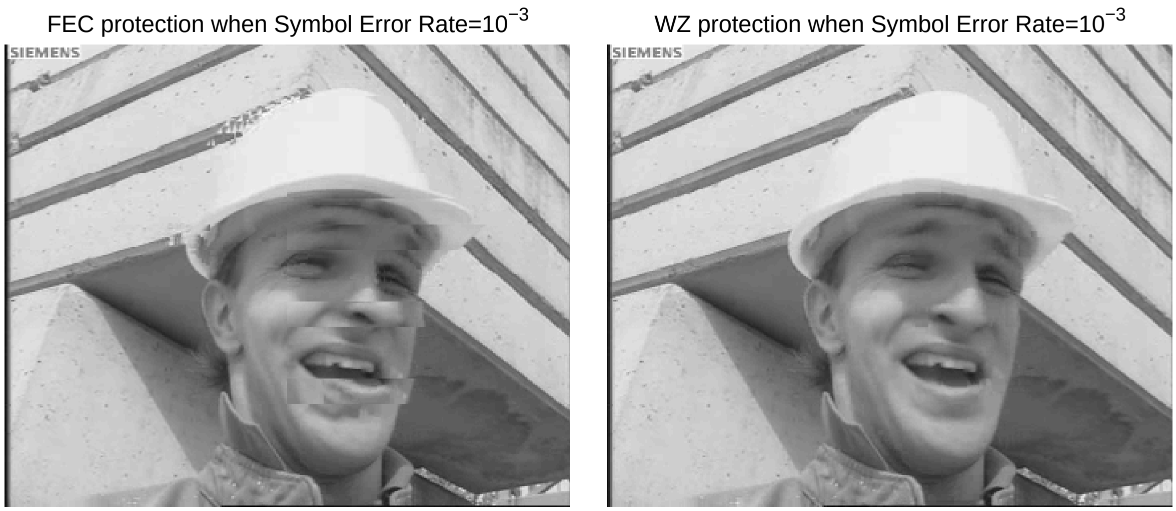

Figure 5. Wyner-Ziv coding gives superior video quality with imperceptible quantization artifacts at high symbol error rates. The figure compares the performance of our scheme with traditional FEC at a symbol error rate of 10^{-3}.

Fig. 6 shows similar results for transmission of the Coastguard CIF sequence encoded at 3.175 Mbps (QP=10). For traditional FEC, a (40,36) RS code is applied to the MPEG bitstream, corresponding to additional parity information of 350 Kbps. We show two cases of using forward error protection, each with only 220 Kbps worth of parity information. In the first case, the coarse description is encoded at 1 Mbps (QP=17), and a (44,36) RS code is applied to it, resulting in 220 Kbps of parity information. In the second case, we trade off the quantizer coarseness in favour of even stronger Wyner-Ziv protection. The second case uses a coarse description coded at only 650 Kbps (QP=19), while a (48,36) RS code is applied to it, again resulting in about 220 Kbps of parity information. As seen in Fig. 6, our scheme outperforms FEC, in the sense of achieving acceptable PSNR over a greater range of channel error probabilities.

5. CONCLUSIONS AND ONGOING WORK

In this paper we propose an error-resilient digital video broadcasting scheme using Wyner-Ziv coding concepts. 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 determined by the quantizer coarseness in the Wyner-Ziv description. Since it allows for some distortion in the case of channel errors, the above scheme can 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.

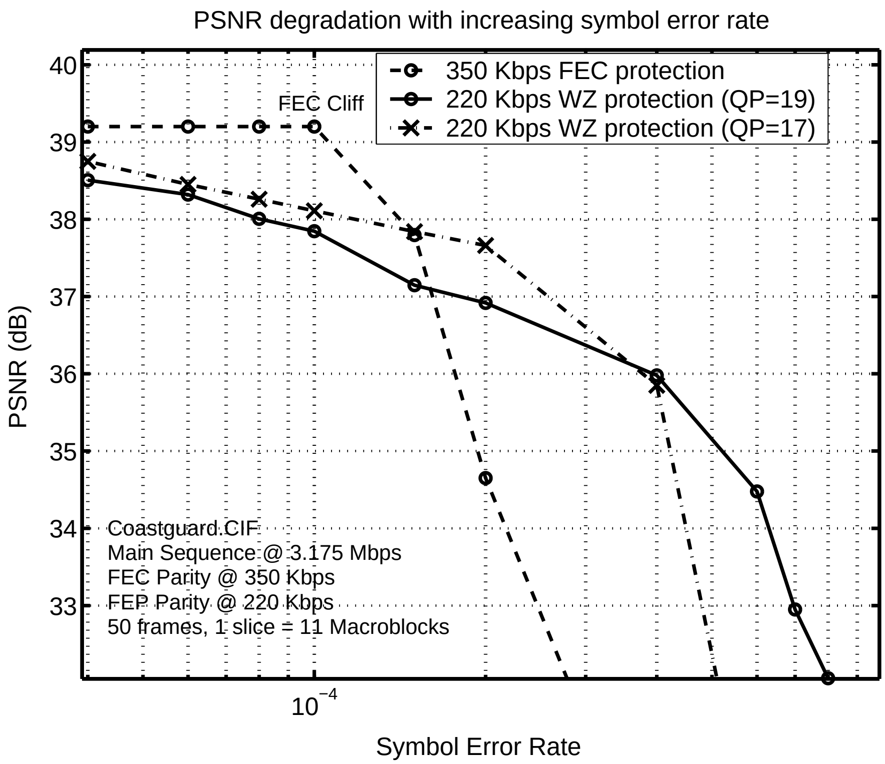

Figure 6. The distortion in the coarse video description can be traded off with the amount of Wyner-Ziv protection to achieve superior overall performance compared to FEC.

The trade-off between the Wyner-Ziv bit-rate and the residual distortion from transmission errors rate (Fig. 6) 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. This is a significant advantage over layered video coding schemes which achieve graceful degradation but incur a high rate-distortion penalty. We recently presented encouraging first results using an embedded pixel-domain Wyner-Ziv codec [12]. Developing such an embedded Wyner-Ziv codec for the more advanced system described in this paper, is the focus of our current efforts. We are also aiming to reduce the complexity of the Wyner-Ziv decoder.

6. ACKNOWLEDGMENTS

This work is supported in part by a National Science Foundation Grant Number CCR-0310376, and a C.V. Starr Southeast Asian Fellowship.

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* **IT-22**, pp. 1–10, Jan. 1976.
2. A. Wyner, "On source coding with side information at the decoder," *IEEE Transactions on Information Theory* **IT-21**, 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* **38**, pp. 60–80, 1978.
4. R. Zamir, "The rate loss in the Wyner-Ziv problem," *IEEE Transactions on Information Theory* **42**, pp. 2073–2084, Nov. 1996.

5. S. Shamai, S. Verdu, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Transactions on Information Theory* **44**, 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*, pp. 63–72, (Snowbird, Utah), Mar. 2001.
7. A. Albanese, J. Blomer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoding transmission," *IEEE Transactions on Information Theory* **42**, 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* **11**, 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* **18**, 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* **15**, 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*, pp. 93–102, (Snowbird, Utah), Mar. 2003.
12. A. Aaron, S. Rane, D. Rebollo, and B. Girod, "Systematic Lossy Source Channel Coding of Video Waveforms," in *Proc. International Conference on Image Processing, Invited Paper*, (Barcelona, Spain), Sept. 2003.
13. R. Puri and K. Ramchandran, "PRISM, A Robust Video Coding Architecture based on Distributed Coding Principles ," in *40th Allerton Conference on Communication*, (Allerton, Illinois), Oct. 2002.
14. A. Sehgal, A. Jagmohan, and N. Ahuja, "A State-Free Causal Video Encoding Paradigm," in *Proc. International Conference on Image Processing, Special Session on Distributed Source Coding*, (Barcelona, Spain), Sept. 2003.
15. A. Jagmohan, A. Sehgal, and N. Ahuja, "Predictive encoding using coset codes," in *Proc. International Conference on Image Processing*, **2**, pp. 29–32, (Rochester, New York), Sept. 2002.
16. 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.
17. Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-C167 (T. Wiegand, ed.), "Committee draft number 1, revision 0 (CD-1)," *ITU-T Recommendation H.26L* , May 2002.
18. T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard," *IEEE Transactions on Circuits and Systems for Video Technology* **13**, pp. 560–576, July 2003.