# ERROR-RESILIENT VIDEO TRANSMISSION USING MULTIPLE EMBEDDED WYNER-ZIV DESCRIPTIONS

*Shantanu Rane, Anne Aaron and Bernd Girod*

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

**ABSTRACT**

This paper proposes systematic lossy error protection of video waveforms using multiple embedded Wyner-Ziv video descriptions. A video signal transmitted over an error-prone channel, without channel coding, constitutes the systematic portion of the transmission. Error protection is achieved by additionally transmitting two or more bitstreams generated by Wyner-Ziv coding of the video signal. The Wyner-Ziv bitstreams contain multiple embedded coarsely quantized descriptions of the original video signal. In the event of channel errors, the Wyner-Ziv descriptions are decoded, thereby limiting the maximum distortion that can occur. The available Wyner-Ziv bitrate is allocated among the multiple embedded descriptions, using the observation that the visibility of slice losses and quantization artifacts depends on the employed coding structure. Experimental results show that the trade-off between error-resilience and residual quantization distortion can be better exploited using multiple embedded Wyner-Ziv descriptions, compared to earlier results with a single Wyner-Ziv bitstream. The system delivers gracefully deteriorating video quality without requiring a layered video compression scheme in the systematic portion of the transmission.

## 1. INTRODUCTION

We are interested in video transmission systems which transmit a bitstream to one or more receivers over an error-prone channel, such as a broadcast channel, or an ad hoc wireless network. For error-resilience, the source bitstream is generally protected using some form of forward error correction (FEC). FEC, along with decoder-based error concealment ensures that the video quality is acceptable for the given application. However, if the channel degrades rapidly owing to fading or shadowing, or if the estimated probability of transmission errors is lower than the actual value, then the FEC parity information is not sufficient for error correction. Hence the video quality degrades rapidly, leading to the undesirable "cliff" effect. It has been shown [1, 2, 3], that using a Wyner-Ziv bitstream in a systematic lossy error protection framework provides improved error-resilience compared to FEC. The Wyner-Ziv bitstream contains a coarsely quantized video description which is decoded when the original video bitstream is corrupted by channel errors, ensuring that the video quality degrades slowly as the probability of transmission errors increases. This paper proposes the use of multiple Wyner-Ziv bitstreams containing embedded video descriptions, so that the trade-off between error-resilience and the residual distortion from coarse quantization in the Wyner-Ziv codec can be exploited more meaningfully.

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 [4, 5, 6]. It was proved that the minimum encoding rate for a source sequence $X$, at 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. It has been shown in [7] that, for mean-squared error distortion, the rate loss associated with ignoring the side information at the encoder is bounded from above, and is actually zero for jointly Gaussian $X$ and $Y$.

The Wyner-Ziv problem is closely related to the problem of systematic lossy source-channel coding [8]. In this configuration, a 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 this configuration in [8].

The systematic coding framework was used by Pradhan and Ramchandran [9] for enhancing the quality of images corrupted by additive white Gaussian noise, using digital side information. In the prequel to this work, we presented a practical systematic lossy error protection (SLEP) system for error-resilient video broadcasting [10, 3, 2]. The Wyner-Ziv codec in the SLEP scheme consists of a hybrid video codec and a Reed-Solomon Slepian-Wolf codec.[1].

The undesirable FEC "cliff" effect can also be mitigated by using Priority Encoding Transmission (PET) [11] which assigns unequal FEC protection to different parts of the video bitstream depending upon their relative importance. This approach, which ensures graceful degradation of image quality in the presence of channel errors, has been exploited by layered video coding schemes [12, 13, 14]. 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 decoded video quality *without the need for a layered representation* in the systematic portion of the transmission. A comparison of the error-resilience performance of the SLEP scheme and that of layered coding with unequal error protection has been presented in [15].

The remainder of this paper is organized as follows. In Section 2, we explain the principle of systematic lossy error protection, and its extension to multiple embedded Wyner-Ziv streams. We then describe a practical implementation of the embedded Wyner-Ziv codec. In Section 3, we present experimental results which show the advantages of embedded Wyner-Ziv coding for digital video broadcasting.

## 2. SYSTEMATIC LOSSY ERROR PROTECTION SCHEME

### 2.1. Wyner-Ziv Coding for Error-Resilient Video Transmission

The concept of systematic lossy error protection (SLEP) is illustrated in Fig. 1, using MPEG video compression as an example. At the

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

[1]In [10, 3], our scheme is referred to as forward error protection (FEP), which can be confused with forward error correction (FEC). In subsequent work, we use the term systematic lossy error protection (SLEP), to prevent any confusion.

0-7803-9134-9/05/$20.00 ©2005 IEEE

transmitter, the input video signal $S$ is compressed independently by an MPEG video coder and a Wyner-Ziv coder. 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, which are corrected, up to a certain residual distortion, by the Wyner-Ziv decoder. The Wyner-Ziv code can be thought of as an independent description of the input video $S$, but with coarser quantization. To mitigate 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

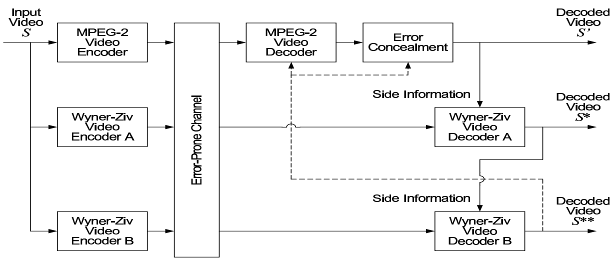


**Fig. 1**. Wyner-Ziv decoder uses decoded error-concealed video waveform as side information in the systematic lossy source-channel coding setup. With an embedded Wyner-Ziv codec, graceful degradation of video quality is obtained without a layered video representation.

bits must be sent to allow error-free reconstruction of the coarser second description, employing the decoded video $S'$ as side information. The Wyner-Ziv description and the 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 Wyner-Ziv representation transmitted at very low bit-rate in the Wyner-Ziv bitstream limits the maximum degradation that can occur. The trade-off between distortion due to transmission errors and Wyner-Ziv bit-rate 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. Such a system is shown in Fig. 1 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, 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 using the decoded output symbols from Wyner-Ziv decoder 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 video coding scheme.

### 2.2. Wyner-Ziv Codec

We now briefly describe a practical Wyner-Ziv codec constructed from well-understood components viz., quantizers, entropy coders, and Reed-Solomon (RS) codecs. First, a Wyner-Ziv bitstream corresponding to a coarsely quantized video description is generated using the method in [3]. 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 constitute the Wyner-Ziv bitstream. *The systematic portion of the RS encoder output is discarded.*

When transmission errors occur, the decoder generates a coarsely quantized version of the received prediction error signal. This coarse version is an error-prone copy of the Wyner-Ziv description, which serves as side information for the RS decoder. Using the parity symbols and error-prone Wyner-Ziv description, the RS decoder performs erasure decoding to obtain the error-free Wyner-Ziv description. A fallback mechanism substitutes the lost slices in the main video sequence with their coarsely quantized counterparts.

To exploit the trade-off between coding efficiency and error-resilience, a video bitstream typically consists of I slices (most significant, highest bit-rate), P slices and B slices (least significant, smallest bit-rate). For improved error-resilience this coding structure must be taken into account when the Wyner-Ziv bitstream is constructed, i.e., The RS Slepian-Wolf coder must output varying amounts of parity symbols for I, P, and B slices. For a video slice of length $l$ symbols, at a symbol error probability $p$ of a memoryless channel, the probability that the slice is corrupted is given by $1-(1-p)^l \approx lp$ for small $p$, i.e., the probability of losing a slice is proportional to its length. Thus, an I slice is $s_I = L_I/L_B$ times more likely to be lost, than a B slice, where $L_I, L_B$ are the average lengths of I and B slices in the main video description. Therefore, for every one parity slice appended to a B frame, our system appends $s_I$ parity slices to the I frame at the beginning of the GOP. Similarly, for every one parity slice appended to a B frame, we append $s_P = L_P/L_B$ parity slices to each P frame in the GOP. Let $l_I, l_P, l_B$ be the average lengths of I, P, and B slices in the Wyner-Ziv description and let $m_I, m_P, m_B$ be the number of I, P, and B frames in one GOP. Based on the priorities assigned above, let the number of parity slices for the I, P, and B frames be $s_I x, s_P x, s_B x$, where $s_B = L_B/L_B = 1$. Then, since only the parity slices are transmitted in the Wyner-Ziv bitstream, the Wyner-Ziv bit-rate is given by $R_{WZ} = (m_I l_I s_I + m_P l_P s_P + m_B l_B s_B)x$, which can be solved for $x$, because all other quantities are known. In this way, unequal Wyner-Ziv protection is assigned within a single Wyner-Ziv description.

### 2.3. Embedded Wyner-Ziv Codec

Now consider the generation of a second Wyner-Ziv bitstream, which contains a video description with finer quantization than that described in the preceding section. For this, the difference between the original transformed prediction error and the coarsely quantized Wyner-Ziv description is obtained and then finely quantized and entropy-coded, as shown in Fig. 2. The resulting bitstream is input to a RS Slepian-Wolf encoder which applies RS coding across the video slices, and transmits only the parity symbols. This method of generating the embedded Wyner-Ziv description is reminiscent of SNR-scalable video coding, with the important distinction that only the parity bitstreams corresponding to the Wyner-Ziv descriptions are transmitted, while the systematic portions are regenerated (possibly with errors) at the decoder. Note that, just like SNR-scalable video

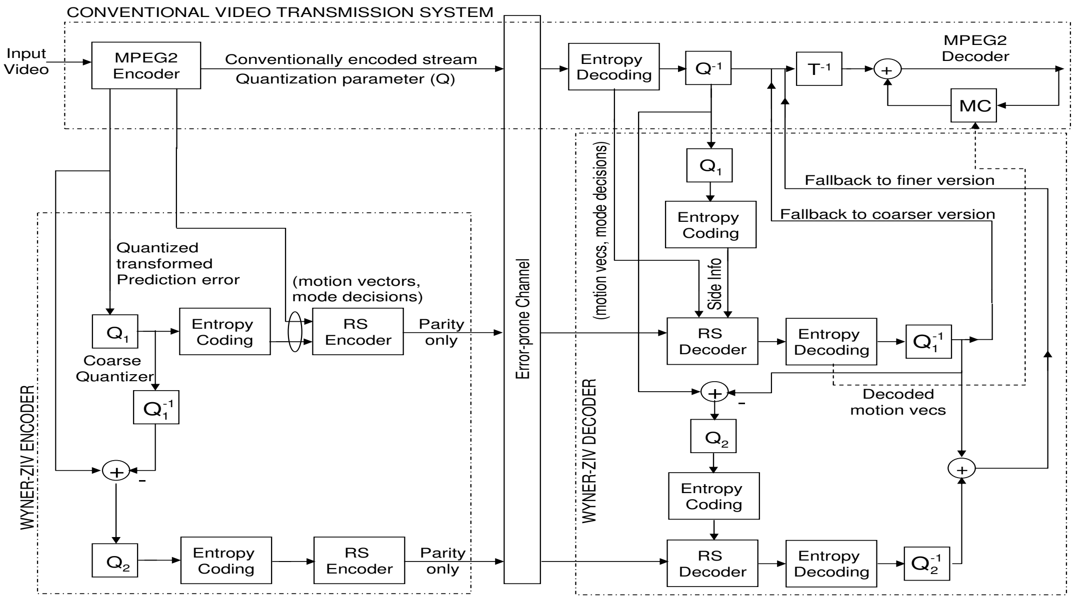


**Fig. 2**. Implementation of systematic lossy error protection by combining MPEG coding and Reed-Solomon codes across slices.

coding, the decoder can only recover the second finer Wyner-Ziv description if the first coarser Wyner-Ziv description has been successfully decoded, and not otherwise. Now consider allocation of the available bit-rate among the finer Wyner-Ziv description and the embedded coarser Wyner-Ziv description. Clearly, a larger share of the Wyner-Ziv bit-rate must be allocated to the more significant coarser Wyner-Ziv description. This is done in such a way that the number of R-S parity slices for a given frame-type, is mid-way between the number of parity slices possible if only one of the two descriptions were available, and Wyner-Ziv protection was carried out as in Section 2.2. For e.g., from Section 2.2, if $n_P^c$ parity slices were used for P frame of the coarse description alone, and $n_P^f$ parity slices were used for P frames of the fine description alone, then in the embedded scheme, $(n_P^c + n_P^f)/2$ parity slices are used for P frames of the embedded coarse Wyner-Ziv description. The remaining bit-rate is allocated to the finer Wyner-Ziv description. Experimentally, this ad hoc approach provided the best resilience-quality trade-off.

## 3. EXPERIMENTAL RESULTS

We now describe the results of applying the SLEP system to error-resilient MPEG-2 video broadcasting. In our experiment the systematic transmission consists of the *Foreman.CIF* sequence encoded at 2 Mbps. For error-resilience, an additional bit-rate of 222 Kbps is available, i.e., for the system of Fig. 2, the sum of the parity bit-rates transmitted by the RS Slepian-Wolf encoders is 222 Kbps.

### 3.1. Unequal Wyner-Ziv protection within one SLEP bitstream

First we consider the advantages of exploiting the coding structure of the video bitstream, as described in Section 2.2. In Fig. 3, the dashed curves indicate the variation in PSNR for the system in [3], in which the I, P, and B slices are not treated differently. The solid curves describe the performance of the proposed system. The bit-rate for error-resilience is the same in each case. When only a 500 Kbps Wyner-Ziv description is available, the RS Slepian-Wolf codes for I, P and B frames are (36,18), (29,18), and (22,18) respectively. When only a 1 Mbps Wyner-Ziv description is available, they are (27,18), (23,18), and (20,18) respectively. As expected, protecting the I, P and B slices after taking into account their lengths and the number of times they occur in the sequence, yields superior error-resilience.

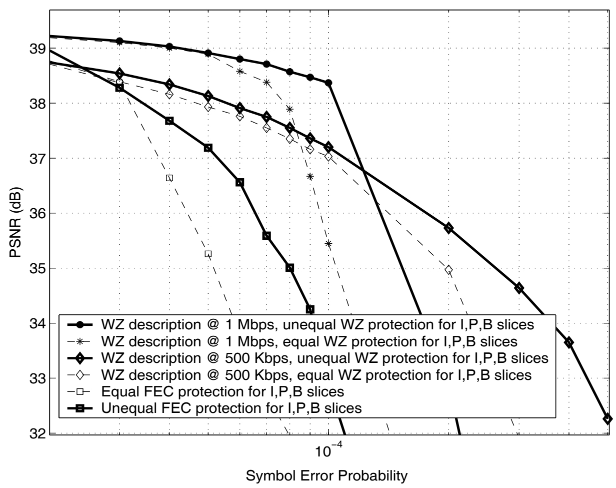


**Fig. 3**. Error resilience improves when unequal Wyner-Ziv protection is assigned to the I, P, B frames in a single Wyner-Ziv description. The transmitted Wyner-Ziv bit-rate is 222 Kbps for each curve.

### 3.2. Embedded Wyner-Ziv Coding

In the second part of the experiment, we allocate the available 222 Kbps bit-rate among two Wyner-Ziv streams, which are generated according to the procedure described in Section 2.3. In the experiment, the first coarser Wyner-Ziv description has a source coding bit-rate of 500 Kbps, and a transmitted Wyner-Ziv bit-rate of 166 Kbps. The second finer Wyner-Ziv description has a source coding bit-rate of 1 Mbps, and a transmitted Wyner-Ziv bit-rate of 56 Kbps. Within each Wyner-Ziv bitstream, the rates of the RS Slepian-Wolf codes for the I, P and B frames are decided according to the procedure described in Section 2.2. The RS codes used are (32,18), (26,18), (21,18) for

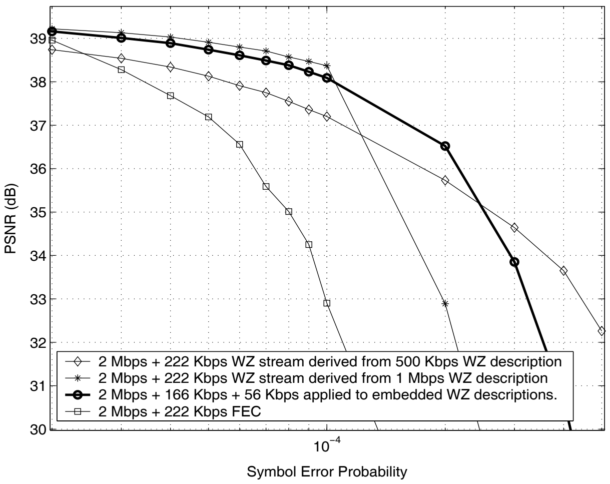


**Fig. 4**. To achieve graceful degradation of video quality, a coarse Wyner-Ziv description encoded at 500 Kbps is embedded inside a finer Wyner-Ziv description encoded at 1 Mbps. The available error-resilience bit-rate of 222 Kbps is then shared among the two descriptions.

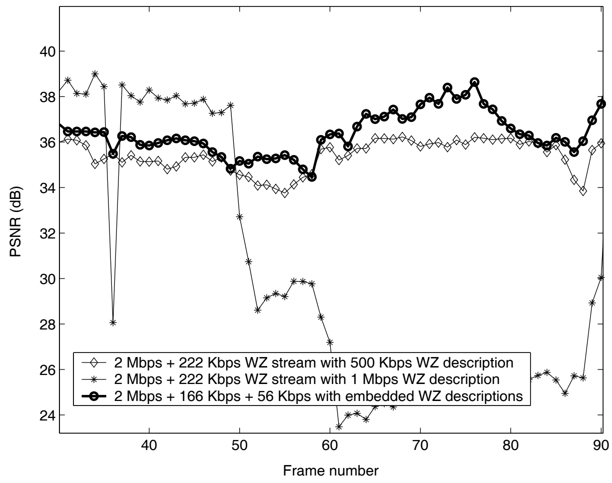


**Fig. 5**. At an error probability of $2 \times 10^{-4}$, SLEP with a finely quantized Wyner-Ziv description alone cannot provide adequate viewing quality because of decoder failure and error propagation, while SLEP with the coarsely quantized Wyner-Ziv description alone incurs more distortion due to coarse quantization. The embedded Wyner-Ziv scheme, exploits the resilience-quality trade-off better than using either of the two Wyner-Ziv descriptions alone.

the coarser Wyner-Ziv description, and (23,18), (20,18), (19,18) for the finer Wyner-Ziv description. As shown in Fig. 4, at low symbol error rates, the decoded video quality provided by the embedded Wyner-Ziv coding scheme is close to that obtained by using the (finer) 1 Mbps Wyner-Ziv description. At high symbol error rates, the decoded video quality is closer to that obtained using the (coarser) embedded 500 Kbps Wyner-Ziv description. Thus the trade-off between resilience to transmission errors, and the residual distortion resulting from Wyner-Ziv quantization is exploited to obtain better overall performance than using any of the two Wyner-Ziv descriptions alone. Since Fig. 4 contains average PSNR values, it demonstrates the graceful degradation property of embedded Wyner-Ziv coding, but does not show the instantaneous effects of decoder failure. To appreciate the advantage of using embedded Wyner-Ziv coding from the point of view of mitigating error propagation within a video sequence, refer to Fig. 5, which shows the variation of PSNR with time for a simulation trace at a symbol error probability of $2 \times 10^{-4}$.

## 4. CONCLUSIONS

Error-resilient video transmission can be accomplished by using Wyner-Ziv coded bitstreams in a systematic lossy source-channel coding framework. It has been shown experimentally that, even with one Wyner-Ziv bitstream, the error-resilience provided by the SLEP system can be improved by taking into account the coding structure of the video bitstream and assigning unequal Wyner-Ziv protection to I, P and B frames. An embedded Wyner-Ziv coding scheme has been described, in which two or more Wyner-Ziv bitstreams are transmitted, corresponding to two embedded Wyner-Ziv video descriptions. By appropriately allocating the bit-rates between the chosen Wyner-Ziv descriptions, and within each Wyner-Ziv description, the trade-off between error-resilience and residual distortion from transmission errors can be exploited. In this way, graceful degradation of video quality is obtained without using layered (scalable) video coding for the systematic portion of the transmission.

## 5. REFERENCES


[1] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proc. IEEE, Special Issue on Advances in Video Coding and Delivery*, vol. 93, no. 1, pp. 71–83, Jan. 2005.

[2] S. Rane and B. Girod, "Analysis of Error-Resilient Video Transmission based on Systematic Source-Channel Coding," in *Picture Coding Symposium (PCS 2004)*, San Francisco, CA, Dec. 2004.

[3] S. Rane, A. Aaron, and B. Girod, "Systematic lossy forward error protection for error resilient digital video broadcasting - A Wyner-Ziv coding approach," in *Proc. IEEE International Conference on Image Processing*, Singapore, Oct. 2004.

[4] 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.

[5] 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.

[6] 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.

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

[8] 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.

[9] 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.

[10] 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.

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

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

[13] 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.

[14] 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.

[15] S. Rane and B. Girod, "Systematic lossy error protection versus layered coding with unequal error protection," in *Proc. SPIE Image and Video Communications and Processing*, San Jose, CA, Jan. 2005.