

Systematic Lossy Error Protection of Video based on H.264/AVC Redundant Slices

Shantanu Rane* and Bernd Girod†
Information Systems Laboratory
Stanford University, Stanford, CA 94305.
{srane,bgirod}@stanford.edu

Abstract

We propose the use of H.264 redundant slices for Systematic Lossy Error Protection (SLEP) of a video signal transmitted over an error-prone channel. In SLEP, the video signal is transmitted to the decoder without channel coding. Additionally, a Wyner-Ziv encoded version of the video signal is transmitted in order to provide error-resilience. In the event of channel errors, the Wyner-Ziv description is decoded as a substitute for the error-prone portions of the primary video signal. Since the Wyner-Ziv description is typically coarser than the primary video signal, SLEP is a *lossy* error protection technique which trades-off residual quantization distortion for improved error-resilience properties, such as graceful degradation of decoder picture quality. We describe how H.264 redundant slices can be used to generate the Wyner-Ziv description, and present simulation results to demonstrate the advantages of this method over traditional methods such as FEC.

Keywords: Systematic lossy source-channel coding, Systematic Lossy Error Protection, Wyner-Ziv coding, Slepian-Wolf coding, side information, redundant slices, error-resilient video coding, distributed video coding.

1. INTRODUCTION

In a video transmission system, a video bitstream is transmitted over an error-prone channel such as a TV broadcast channel, or a wireless ad-hoc network, or over the internet. A variety of methods can be used to make the bitstream resilient to channel errors. Systematic Lossy Error Protection (SLEP)¹⁻³ is a recently proposed error-resilience technique based on lossy source channel coding, which trades off decoded picture quality in favor of improved error resilience. Owing to this tradeoff, SLEP provides gracefully deteriorating video quality under worsening channel conditions, and thus mitigates the “cliff effect” encountered by systems using conventional forward error correction (FEC). In this paper, we are interested in implementing a SLEP system for H.264 video transmission. The structure of the SLEP scheme is such that it can be designed easily using redundant slices, one of the error resilience tools provided within the H.264/AVC standard.

A redundant slice is an alternative representation of a video slice, which may use different quantization parameters, different reference pictures, different mode decisions, and different motion vectors than those used in the encoding of the primary slice. If the primary slice is received correctly, the redundant slice is discarded. However, if the primary slice is received in error, the redundant slice can be decoded in order to limit the distortion caused by the error in the primary bitstream. Typically, the parameters for encoding the redundant slice should be selected in such a way that there is no visual discrepancy between the primary and redundant slices. The concept of redundant slices is an extension of the redundant picture concept used for video redundancy coding⁴ in H.263+. Even though the redundant slices have been described as a feature in the H.264 documentation⁵ and related publications,^{6,7} no actual results have been presented. The proposal of “key pictures”

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

†This work has been supported by NSF Grant No: CCR-0310376.

in multicast streaming⁸ employs redundant slices, but in a trivial way, by restricting the redundant slice to be an identical copy of the primary slice. In this paper, we provide experimental results in which the redundant slices are indeed different from the primary slices and examine the resulting degradation in video quality. We then use redundant slices to implement a SLEP system, the concept of which is described in the next section.

2. SYSTEMATIC LOSSY ERROR PROTECTION

Systematic Lossy Error Protection (SLEP) is based on the theoretical framework of systematic lossy source-channel coding. In this framework, a signal is transmitted over an error-prone analog channel without channel coding. On a separate channel, an encoded version of the signal is transmitted as digital enhancement information. The receiver decodes this enhancement information, using the received error-prone output of the analog channel as side-information. Information theoretic optimality conditions for this lossy source channel coding scenario have been derived by Shamai, Verdú and Zamir⁹. Of particular interest is the idea that the encoding of digital enhancement information falls under the category of lossy source coding with decoder side information, namely the Wyner-Ziv coding problem¹⁰. In deriving the set of achievable rates for this problem,¹¹ Wyner and Ziv showed that the minimum encoding rate for a source sequence given side information at the decoder alone, is lower bounded by the encoding rate given side information at the encoder and the decoder. Interestingly, it was shown that the rate loss resulting from the encoder's lack of knowledge of the side-information is also bounded from above.¹² With the emergence of new applications in low complexity video encoding, and error-resilient video compression, Wyner-Ziv coding and related ideas are attracting significant interest in the video coding community.³

The concept of systematic lossy error protection (SLEP) is illustrated in Fig. 1, using digital video transmission as an example. At the transmitter, the input video signal S is compressed independently by a hybrid video coder and a Wyner-Ziv coder. Since the 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 compressed video signal 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 video bitstream is decoded and transmission errors are concealed, resulting in the decoded video S' . Even after concealment, S' contains 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. 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' .

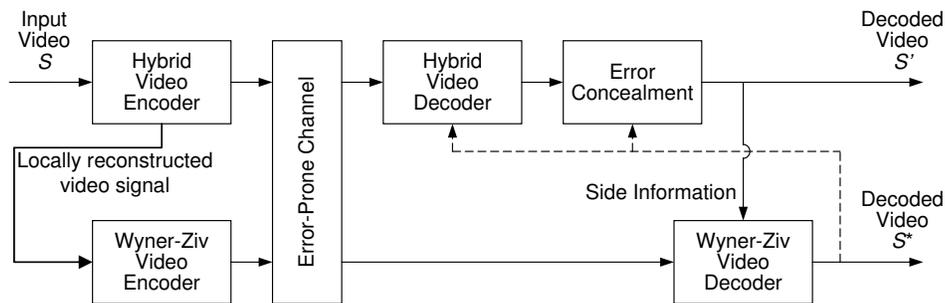


Figure 1. Wyner-Ziv decoder uses decoded error-concealed video waveform as side information in systematic lossy source-channel setup.

When transmission errors occur, 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 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 bit-rate in the Wyner-Ziv bitstream limits the maximum degradation that can occur.

3. USE OF H.264 REDUNDANT SLICES FOR SLEP

A Wyner-Ziv encoder can be conceptually imagined as a coarse quantizer followed by a Slepian-Wolf encoder. A Slepian-Wolf encoder¹³ performs lossless encoding of source data assuming that the decoder has access to some side information which is not known to the encoder. Thus, a Wyner-Ziv bitstream can be generated in two stages: (1) Generate a coarsely quantized version of the original video signal, and (2) Perform Slepian-Wolf coding of the coarse description. The first stage, i.e., generating a second video description, can be readily accomplished by turning on the redundant slice feature in the encoder. For the second stage, we note that a Slepian-Wolf encoder essentially tries to protect the input signal from errors introduced by a virtual “correlation” channel between the input and the side-information. Thus Slepian-Wolf coding is closely related to channel coding. To implement a Slepian-Wolf code, a Reed-Solomon (RS) code with byte-long symbols is applied across the redundant slices generated by the H.264 encoder, and *only the parity slices* generated by the RS code are sent to the decoder. A Wyner-Ziv encoder constructed in this manner, using redundant slices and Reed-Solomon Slepian-Wolf encoding is shown in Fig. 2.

In the event of channel errors, a redundant description of the error-prone frame is regenerated at the decoder. The Wyner-Ziv decoder now possesses an error-prone redundant bitstream, in which locations of the error-prone slices are known. A Reed-Solomon decoder can combine these error-prone slices with the parity slices received in the Wyner-Ziv bitstream and perform erasure decoding across the redundant slices. In the language of distributed video coding, the RS decoder acts as a Slepian Wolf decoder, which decodes the parity slices using the error-prone redundant slices as side information. Upon erasure decoding, we obtain the correct redundant versions of the slices which were lost from the main description. Finally, these redundant slices are decoded and displayed in lieu of the lost slices. The coarse quantization in the redundant slices results in a quantization mismatch which propagates to the succeeding frames, but this scheme avoids severe reduction in picture quality resulting from slice loss or insufficient error concealment.

The Wyner-Ziv decoding algorithm can be greatly simplified, if the Wyner-Ziv bitstream uses the same mode-decisions and motion vectors as those used in the primary bitstream. This means that the redundant slices differ from the primary slices only in the quantization parameter \ddagger . With this modification, regeneration of the redundant description at the decoder is reduced to requantization of the received prediction error signal followed by entropy coding, as shown in Fig. 2. Analysis and optimization of such a system (using MPEG-2 encoded video) has been carried out in our previous work.¹⁴

[‡]Strictly speaking, this is a slight departure from the SLEP concept described in the previous section because the Wyner-Ziv codec now operates on the prediction error signal rather than on the video signal itself. Constraining the redundant slices in this way reduces their coding efficiency, but the requantization method is chosen because of its simplicity and lower complexity, compared to performing a full-fledged re-encoding of the decoded error-concealed output video signal, as required by the system in Fig. 1.

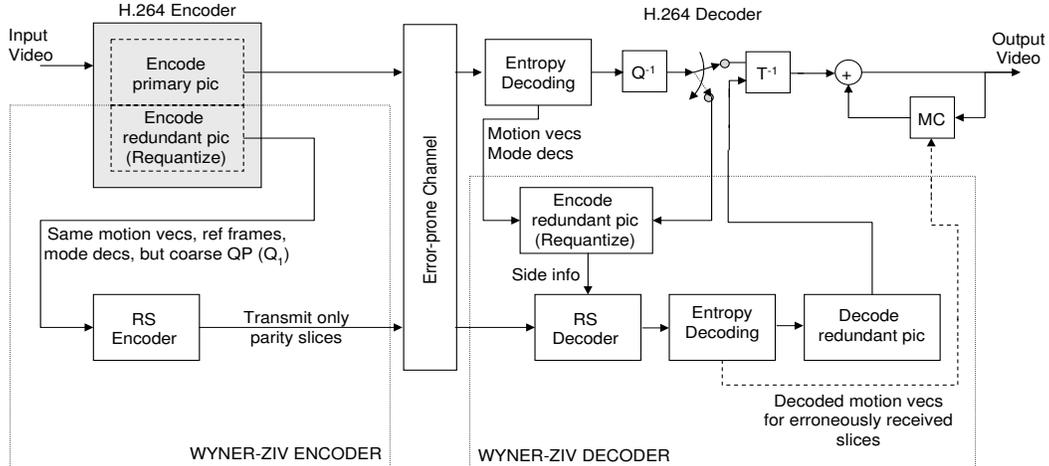


Figure 2. A Wyner-Ziv codec is constructed using H.264 redundant slices in conjunction with a Reed-Solomon coder acting as a Slepian-Wolf coder.

4. EXPERIMENTAL RESULTS

4.1. Error-Resilience by Transmitting Redundant Slices to the Decoder

In this section, we consider transmission of a video sequence over an error-prone channel, and examine the effect of replacing the lost slices by redundant slices. The primary video sequence consists of the first 100 frames of the Foreman.CIF sequence encoded at 408 kbps ($QP=28$). The GOP structure is I-B-P-B-P-.... The video frames are partitioned into slices, such that one row of macroblocks corresponds to one video slice. For error resilience, along with the primary video slices, redundant slices are also sent to the decoder. We control the quantization parameter of the redundant picture, and allow the encoder to choose its own mode decisions and motion vectors, independently of the primary slice. To demonstrate the error-resilience provided by redundant slices, video slices are dropped at random from the primary picture, and are concealed by decoding the corresponding redundant slices. Fig. 3 shows the degradation in picture quality when the percentage of dropped primary picture slices is increased from 2% to 10%. Each data-point in Fig. 3 corresponds to the PSNR averaged over 25 realizations, in which slices are dropped independently with a given error probability. It is evident that there exists a trade-off between the degradation in the picture quality due to the redundant slices and the overhead, in terms of the excess bit-rate required for the redundant slices.

4.2. SLEP using Redundant Slices

By using redundant slices in a systematic lossy source channel coding framework, we now show experimentally that it is possible to obtain the error-resilience properties of the previous section, but at a much lower bit-rate. As explained in Section 3, the redundant slices are now generated using the same reference pictures, mode decisions and motion vectors that were used in the primary picture, with the only difference being that a coarser quantization parameter is chosen. The primary description again consists of the Foreman CIF sequence encoded at 408 kbps. To generate the Wyner-Ziv bit-stream, Reed-Solomon codes are applied across the video slices of the redundant description, and only the parity symbols are transmitted. The error resilience bit-rate is chosen to be as close as possible to 40.8 kbps, which is 10% of the source coding rate of the primary description.

We compare the performance of FEC with 5 SLEP schemes, each using a different redundant description. The competing SLEP schemes use redundant descriptions obtained with $QP = 32$,

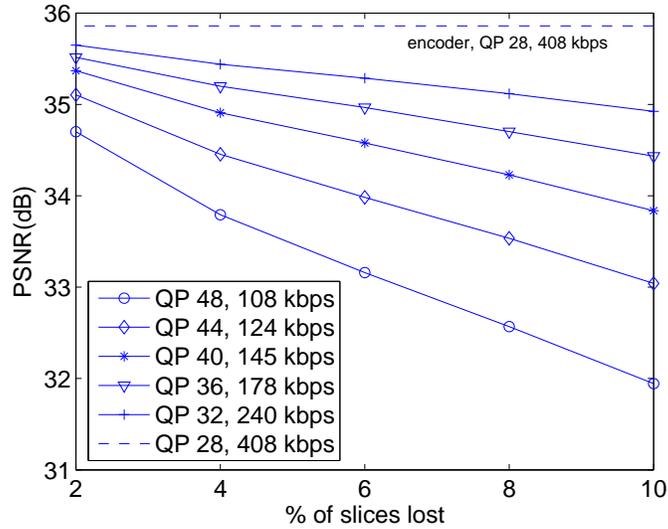
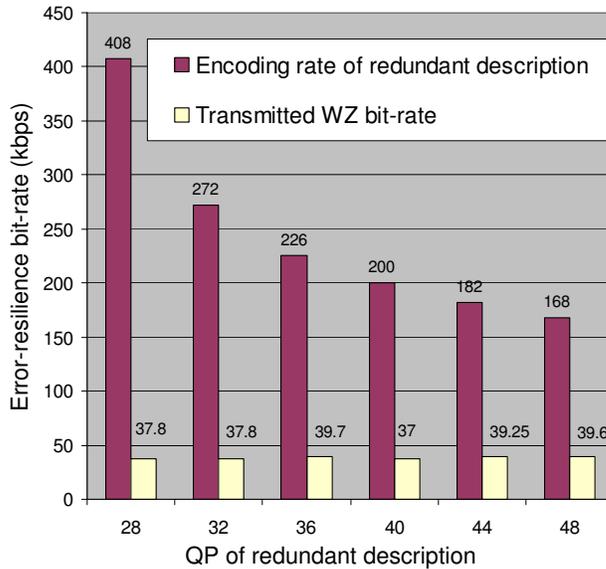


Figure 3. Lost slices are concealed using their redundantly encoded counterparts. The degradation in video quality reduces when the extra bit-rate used to encode the redundant slices is increased.

36, 40, 44, and 48. For the six competing schemes, Fig. 4(a) shows the encoding bit-rates of the redundant descriptions and the corresponding transmitted Wyner-Ziv bit-rates. Note that when $QP = 28$, SLEP reduces to conventional FEC. It is clear that, since the Wyner-Ziv bit-rate is almost constant across all schemes, the SLEP scheme with a lower quality redundant description has stronger error-resilience properties. The table in Fig. 4(b) lists the number of parity slices transmitted in the Wyner-Ziv bit-stream with each I, P, and B picture, for each of the six schemes. We emphasize that the Wyner-Ziv bitstream consists solely of these parity slices, whose bit-rates are given in the rightmost column of the table, and that the redundant descriptions themselves are not sent to the decoder. Upon reception of an error-prone frame, Wyner-Ziv decoding is performed as described in Section 3. If the Wyner-Ziv bit-rate is insufficient to perform Reed-Solomon Slepian-Wolf decoding, the redundant description cannot be regenerated at the decoder, and previous frame error concealment is used to conceal the erroneous slices in the primary description. If Wyner-Ziv decoding is successful, then the erroneous slices are replaced by their regenerated redundant counterparts. The instantaneous effect of independent symbol errors on the video quality is shown in Fig. 5, for a symbol error probability of 0.0005.[§] Also shown in the figure are the frame qualities of the encoder’s reconstruction of the primary video description, i.e., the error-free case, and the frame qualities obtained by using only previous frame error concealment with zero error resilience bit-rate. As seen in the figure, the allotted FEC bit-rate is insufficient for error protection. Depending upon the redundant description used by the Wyner-Ziv codec, we observe a tradeoff between error-resilience and decoded video quality. For e.g., the redundant description encoded with $QP=48$ is the most resilient but loses up to 3.5 dB due to the quantization mismatch incurred during Wyner-Ziv decoding. On the other hand, the redundant description with $QP=40$ cannot be recovered perfectly, since Wyner-Ziv decoding fails for a few B frames, but this description has higher average video quality due to a smaller quantization mismatch. A better appreciation of the subjective quality of the decoded signal can be obtained from Fig. 6, which shows a portion of decoded frame #81, for

[§](1 symbol = 1 byte). Owing to the absence of resynchronization information within a slice, the entire slice is assumed to be erased, if one or more symbols from it are received in error. In terms of channel coding overhead, Reed Solomon codes are optimal for erasure decoding, and this provides the rationale for using a Reed-Solomon code in the Slepian-Wolf codec.



(a) a

QP of redundant description	No. of parity slices transmitted per frame by Reed-Solomon Slepian-Wolf coder			Transmitted Wyner-Ziv bit-rate (kbps)
	I pic	P pic	B pic	
48	6	5	3	39.6
44	6	5	2	39.25
40	6	4	2	37
36	6	4	1	39.7
32	6	3	1	37.8
28 (FEC)	5	2	0	37.8

(b) b

Figure 4. For the price of increased decoding complexity, SLEP reduces the transmitted error-resilience bit-rate by transmitting only the parity symbols corresponding to the redundant descriptions, rather than transmitting the full descriptions. Since the error-resilience bit-rate is fixed to about 10% of the primary description bit-rate, the Wyner-Ziv protection becomes stronger as the redundant description becomes coarser.

each of the six plots in Fig. 5. The subjective comparison not only shows the benefits of lossy error protection over FEC, but also demonstrates the trade-off associated with redundant descriptions of different qualities.

5. CONCLUSIONS AND ONGOING WORK

We proposed a system which implements systematic lossy source-channel coding using redundant slices, a standard-compliant feature in the H.264/AVC specification. The lossy error protection system uses Wyner-Ziv coding, and the Wyner-Ziv video codec is constructed with the redundant slices in conjunction with a Reed-Solomon Slepian-Wolf coder. The resulting SLEP system not only leverages the error-resilience property of redundant slices, but by applying Wyner-Ziv coding concepts, does so at a much lower bit-rate than that achieved by straightforward transmission of the redundant slices. In scenarios where the probability of channel errors is unknown or fluctuating, the SLEP scheme outperforms traditional FEC by providing a more graceful degradation of video quality over a wider range of slice loss probabilities.

The performance of the SLEP scheme can be enhanced by using redundant slices in conjunction with flexible macroblock ordering (FMO) which enables redundant encoding of selected portions of a video frame.¹⁵ Integration of FMO in the SLEP framework and extending the application of SLEP to practical internet video streaming constitutes the focus of our current work.

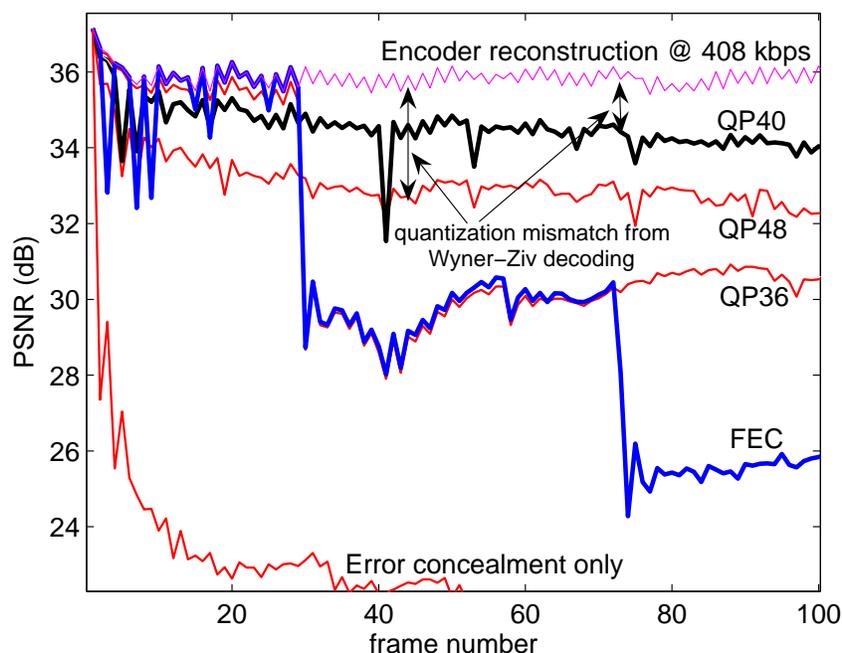


Figure 5. SLEP using redundant slices provides acceptable video quality even when a large number of video slices are lost. At the same bit-rate, FEC protection is insufficient and the video quality degrades rapidly. Selection of the best SLEP description involves a tradeoff between error-resilience of the redundant description and the quantization mismatch resulting from Wyner-Ziv decoding.

REFERENCES

1. 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.
2. S. Rane, A. Aaron, and B. Girod, "Error Resilient Video Transmission using Multiple Embedded Wyner-Ziv Descriptions," in *Proc. IEEE International Conference on Image Processing*, (Genoa, Italy), Sept. 2005.
3. B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proc. IEEE, Special Issue on Advances in Video Coding and Delivery* **93**, pp. 71–83, Jan. 2005.
4. S. Wenger, "Video Redundancy Coding in H.263+," in *Proc. International Workshop on Audio-Visual Services over Packet Networks*, (Aberdeen, UK), Sept. 1997.
5. ITU-T Video Coding Experts Group and ISO/IEC Moving Picture Experts Group, "Study of Final Committee Draft of Joint Video Specification, (ITU-T Rec. H.264, ISO/IEC 14496-10 AVC)," Mar. 2003.
6. J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC: Tools, Performance, and Complexity," *IEEE Circuits and Systems Magazine* **1**, pp. 7–28, Apr. 2004.
7. T. Stockhammer, M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Trans. Circuits and Systems for Video Tech.* **13**, pp. 657–673, July 2003. Special Issue on the H.264/AVC Video Coding Standard.



(a) Encoder reconstruction, PSNR=35.7 dB



(b) Concealment only, PSNR=20.9 dB



(c) FEC, PSNR=25.5 dB



(d) SLEP with redundant QP=36, PSNR=30.9 dB



(e) SLEP with redundant QP=48, PSNR=32.9 dB



(f) SLEP with redundant QP=40, PSNR=34.2 dB

Figure 6. Examination of the decoded picture quality indicates the superior error-resilience of SLEP compared to FEC, and shows the resilience-quality tradeoffs resulting from the choice of the redundant description.

8. Y. K. Wang, M. Hannuksela, and M. Gabbouj, "Error resilient video coding using unequally protected key pictures," in *Proc. International Workshop on Very Low Bit-Rate Video Coding (VLBV 2003)*, Sept. 2003.
9. S. Shamai, S. Verdú, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Transactions on Information Theory* **44**, pp. 564–579, Mar. 1998.
10. A. Wyner, "On source coding with side information at the decoder," *IEEE Transactions on Information Theory* **IT-21**, pp. 294–300, May 1975.
11. A. D. 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.
12. R. Zamir, "The rate loss in the Wyner-Ziv problem," *IEEE Transactions on Information Theory* **42**, pp. 2073–2084, Nov. 1996.
13. J. D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory* **IT-19**, pp. 471–480, July 1973.
14. 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.
15. P. Baccichet and A. Chimienti, "Forward Selective Protection Exploiting Redundant Slices and FMO in H.264/AVC," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006)*, (Toulouse, France), Oct. 2005. Submitted.