# Distributed Video Coding

BERND GIROD, FELLOW, IEEE, ANNE MARGOT AARON,
SHANTANU RANE, STUDENT MEMBER, IEEE, AND DAVID REBOLLO-MONEDERO

*Invited Paper*

*Distributed coding is a new paradigm for video compression, based on Slepian and Wolf's and Wyner and Ziv's information-theoretic results from the 1970s. This paper reviews the recent development of practical distributed video coding schemes. Wyner–Ziv coding, i.e., lossy compression with receiver side information, enables low-complexity video encoding where the bulk of the computation is shifted to the decoder. Since the interframe dependence of the video sequence is exploited only at the decoder, an intraframe encoder can be combined with an interframe decoder. The rate-distortion performance is superior to conventional intraframe coding, but there is still a gap relative to conventional motion-compensated interframe coding. Wyner–Ziv coding is naturally robust against transmission errors and can be used for joint source-channel coding. A Wyner–Ziv MPEG encoder that protects the video waveform rather than the compressed bit stream achieves graceful degradation under deteriorating channel conditions without a layered signal representation.*

***Keywords**—Distributed source coding, low-complexity video coding, robust video transmission, video coding, Wyner-Ziv coding.*

## I. INTRODUCTION

In video coding, as standardized by MPEG or the ITU-T H.26x recommendations, the encoder exploits the statistics of the source signal. This principle seems so fundamental that it is rarely questioned. However, efficient compression can also be achieved by exploiting source statistics—partially or wholly—at the decoder only. This surprising insight is the consequence of information-theoretic bounds established in the 1970s by Slepian and Wolf for distributed lossless coding, and by Wyner and Ziv for lossy coding with decoder side information. Schemes that build upon these theorems are generally referred to as distributed coding algorithms.

The recent first steps toward practical distributed coding algorithms for video have lead to considerable excitement
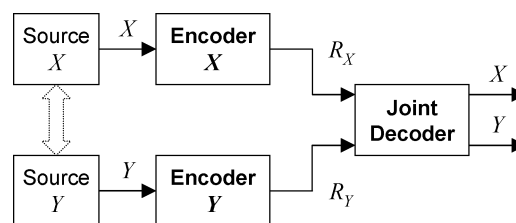
**Fig. 1.** Distributed compression of two statistically dependent random processes $X$ and $Y$. The decoder jointly decodes $X$ and $Y$ and, thus, may exploit their mutual dependence.

in the community. We review these advances in our paper. Distributed coding is a radical departure from conventional, nondistributed coding. Therefore, Section II of our paper will be devoted to understanding the foundations of distributed coding and compression techniques that exploit receiver side information. In Section III, we show how video compression with low encoder complexity is enabled by distributed coding. Distributed coding exploits the source statistics in the decoder and, hence, the encoder can be very simple, at the expense of a more complex decoder. The traditional balance of complex encoder and simple decoder is essentially reversed. Such algorithms hold great promise for new generations of mobile video cameras. In Section IV, we discuss the inherent robustness of distributed coding schemes and then review unequal error protection for video that protects the video waveform (rather than the compressed bit stream) with a lossy source-channel code, another new application that builds upon distributed compression algorithms.

## II. FOUNDATIONS OF DISTRIBUTED CODING

### A. Slepian–Wolf Theorem for Lossless Distributed Coding

Distributed compression refers to the coding of two (or more) dependent random sequences, but with the special twist that a separate encoder is used for each (Fig. 1). Each encoder sends a separate bit stream to a single decoder which may operate jointly on all incoming bit streams and thus exploit the statistical dependencies.
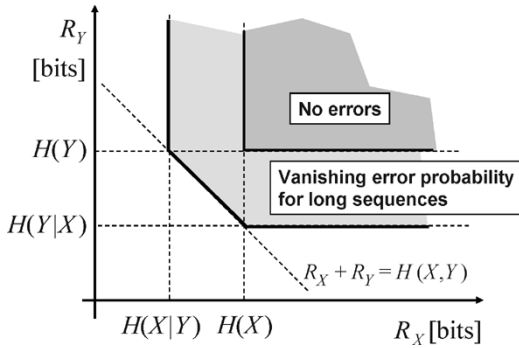
**Fig. 2.** Slepian–Wolf theorem, 1973: achievable rate region for distributed compression of two statistically dependent i.i.d. sources $X$ and $Y$ [1].
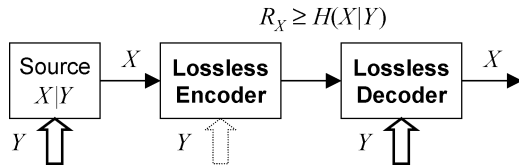
$$R_X \geq H(X|Y)$$



**Fig. 3.** Compression of a sequence of random symbols $X$ using statistically related side information $Y$. We are interested in the distributed case, where $Y$ is only available at the decoder, but not at the encoder.

Consider two statistically dependent independent identically distributed (i.i.d.) finite-alphabet random sequences $X$ and $Y$. With separate conventional entropy encoders and decoders, one can achieve $R_X \geqslant H(X)$ and $R_Y \geqslant H(Y)$, where $H(X)$ and $H(Y)$ are the entropies of $X$ and $Y$, respectively. Interestingly, we can do better with joint decoding (but separate encoding), if we are content with a residual error probability for recovering $X$ and $Y$ that can be made arbitrarily small (but, in general, not zero) for encoding long sequences. In this case, the Slepian–Wolf theorem [1] establishes the rate region (see Fig. 2)

$$R_X + R_Y \geqslant H(X, Y)$$
$$R_X \geqslant H(X|Y), \quad R_Y \geqslant H(Y|X).$$

Surprisingly, the sum of rates $R_X + R_Y$ can achieve the joint entropy $H(X, Y)$, just as for joint encoding of $X$ and $Y$, despite separate encoders for $X$ and $Y$.

Compression with decoder side information (Fig. 3) is a special case of the distributed coding problem (Fig. 1). The source produces a sequence $X$ with statistics that depend on side information $Y$. We are interested in the case where this side information $Y$ is available at the decoder, but not at the encoder. Since $R_Y = H(Y)$ is achievable for (conventionally) encoding $Y$, compression with receiver side information corresponds to one of the corners of the rate region in Fig. 2 and, hence $R_X \geqslant H(X|Y)$, regardless of the encoder's access to side information $Y$.

### B. Practical Slepian–Wolf Coding

Although Slepian and Wolf's theorem dates back to the 1970s, it was only in the last few years that emerging applications have motivated serious attempts at practical techniques. However, it was understood already 30 years ago that

Slepian–Wolf coding is a close kin to channel coding [2]. To appreciate this relationship, consider i.i.d. binary sequences $X$ and $Y$ in Fig. 3. If $X$ and $Y$ are similar, a hypothetical "error sequence" $\Delta = X \oplus Y$ consists of zeros, except for some ones that mark the positions where $X$ and $Y$ differ. To "protect" $X$ against errors $\Delta$, we could apply a systematic channel code and only transmit the resulting parity bits. At the decoder, one would concatenate the parity bits and the side information $Y$ and perform error-correcting decoding. If $X$ and $Y$ are very similar indeed, only few parity bits would have to be sent, and significant compression results. We emphasize that this approach does not perform forward error correction (FEC) to protect against errors introduced by the transmission channel, but instead by a virtual "correlation channel" that captures the statistical dependence of $X$ and the side information $Y$.

In an alternative interpretation, the alphabet of $X$ is divided into cosets and the encoder sends the index of the coset that $X$ belongs to [2]. The receiver decodes by choosing the codeword in that coset that is most probable in light of the side information $Y$. It is easy to see that both interpretations are equivalent. With the parity interpretation, we send a binary row vector $X_p = XP$, where $G = [I|P]$ is the generator matrix of a systematic linear block code $\mathbb{C}_p$. With the coset interpretation, we send the "syndrome" $S = XH$, where $H$ is the parity check matrix of a linear block code $\mathbb{C}_s$. If $P = H$, the transmitted bit streams are identical.

Most distributed source coding techniques today are derived from proven channel coding ideas. The wave of recent work was ushered in 1999 by Pradhan and Ramchandran [3]. Initially, they addressed the asymmetric case of source coding with side information at the decoder for statistically dependent binary and Gaussian sources using scalar and trellis coset constructions. Their later work [4]–[7] considers the symmetric case where $X$ and $Y$ are encoded with the same rate. Wang and Orchard [8] used an embedded trellis code structure for asymmetric coding of Gaussian sources and showed improvements over the results in [3].

Since then, more sophisticated channel coding techniques have been adapted to the distributed source coding problem. These often require iterative decoders, such as Bayesian networks or Viterbi decoders. While the encoders tend to be very simple, the computional load for the decoder, which exploits the source statistics, is much higher. García-Frías and Zhao [9], [10], Bajcsy and Mitran [11], [12], and our own group [13] independently proposed compression schemes where statistically dependent binary sources are compressed using turbo codes. It has been shown that the turbo code-based scheme can be applied to compression of statistically dependent nonbinary symbols [14], [15] and Gaussian sources [13], [16] as well as compression of single sources [10], [17]–[19]. Iterative channel codes can also be used for joint source-channel decoding by including both the statistics of the source and the channel in the decoding process [13], [17]–[21]. Liveris *et al.* [21]–[23], Schonberg *et al.* [24]–[26], and other authors [27]–[30] have suggested that low-density parity-check (LDPC) codes might be a powerful alternative to turbo codes for distributed coding.
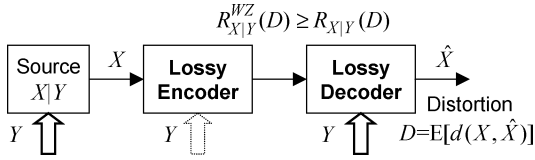
**Fig. 4.** Lossy compression of a sequence $X$ using statistically related side information $Y$.



**Fig. 5.** Practical Wyner–Ziv coder is obtained by cascading a quantizer and a Slepian–Wolf encoder.

With sophisticated turbo codes or LDPC codes, when the code performance approaches the capacity of the correlation channel, the compression performance approaches the Slepian–Wolf bound.

### C. Rate-Distortion Theory for Lossy Compression With Receiver Side Information

Shortly after Slepian and Wolf's seminal paper, Wyner and Ziv [31]–[33] extended this work to establish information-theoretic bounds for lossy compression with side information at the decoder. More precisely, let $X$ and $Y$ represent samples of two i.i.d. random sequences, of possibly infinite alphabets $\mathcal{X}$ and $\mathcal{Y}$, modeling source data and side information, respectively. The source values $X$ are encoded without access to the side information $Y$ (Fig. 4). The decoder, however, has access to $Y$, and obtains a reconstruction $\hat{X}$ of the source values in alphabet $\hat{\mathcal{X}}$. A distortion $D = E[d(X, \hat{X})]$ is acceptable. The Wyner–Ziv rate-distortion function $R_{X|Y}^{WZ}(D)$ then is the achievable lower bound for the bit-rate for a distortion $D$. We denote by $R_{X|Y}(D)$ the rate required if the side information were available at the encoder as well.

Wyner and Ziv proved that unsurprisingly, a rate loss $R_{X|Y}^{WZ}(D) - R_{X|Y}(D) \geqslant 0$ is incurred when the encoder does not have access to the side information. However, they also showed that $R_{X|Y}^{WZ}(D) - R_{X|Y}(D) = 0$ in the case of Gaussian memoryless sources and mean-squared error distortion [31], [33]. This result is the dual of Costa's "dirty paper" theorem for channel coding with sender-only side information [34]–[37]. As Gaussian-quadratic cases, both lend themselves to intuitive sphere-packing interpretations. $R_{X|Y}^{WZ}(D) - R_{X|Y}(D) = 0$ also holds for source sequences $X$ that are the sum of arbitrarily distributed side information $Y$ and independent Gaussian noise [37]. For general statistics and a mean-squared error distortion measure, Zamir [38] proved that the rate loss is less than 0.5 b/sample.

### D. Practical Wyner–Ziv Coding

As with Slepian–Wolf coding, efforts toward practical Wyner–Ziv coding schemes have been undertaken only recently. The first attempts to design quantizers for reconstruction with side information were inspired by the information-theoretic proofs. Zamir and Shamai [39] (see also [40]) proved that under certain circumstances, linear codes and nested lattices may approach the Wyner–Ziv rate-distortion function, in particular if the source data and side information are jointly Gaussian. This idea was further developed and applied by Pradhan *et al.* [3], [41], [42] and Servetto [43], who published heuristic designs and performance analysis focusing on the Gaussian case, based
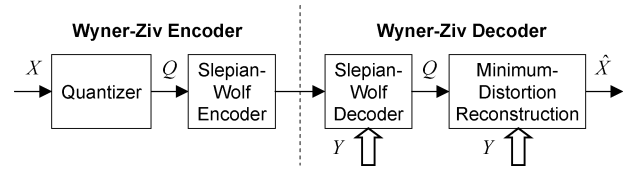
on nested lattices. Xiong *et al.* [44], [45] implemented a Wyner–Ziv encoder as a nested lattice quantizer followed by a Slepian–Wolf coder, and in [46], a trellis-coded quantizer was used instead (see also [47]).

In general, a Wyner–Ziv coder can be thought to consist of a quantizer followed by a Slepian–Wolf encoder, as illustrated in Fig. 5. The quantizer divides the signal space into cells, which, however, may consist of noncontiguous subcells mapped into the same quantizer index $Q$. This setting was considered, e.g., by Fleming, Zhao and Effros [48], who generalized the Lloyd algorithm [49] for locally optimal fixed-rate Wyner–Ziv vector quantization design. Later, Fleming and Effros [50] included rate-distortion optimized vector quantizers in which the rate measure is a function of the quantization index, for example, a codeword length. Unfortunately, vector quantizer dimensionality and entropy code block length are identical in their formulation and, thus, the resulting quantizers either lack in performance or are prohibitively complex. An efficient algorithm for finding globally optimal quantizers among those with contiguous code cells was provided in [51]. Unfortunately, it has been shown that code cell contiguity precludes optimality in general [52]. Cardinal and Van Asche [53] considered Lloyd quantization for ideal Slepian–Wolf coding, without side information.

An independent, more general extension of the Lloyd algorithm appears in our work [54]. A quantizer is designed assuming that an ideal Slepian–Wolf coder is used to encode the quantization index. The introduction of a rate measure that depends on both the quantization index and the side information divorces the dimensionality of the quantizer from the block length of the Slepian–Wolf coder, a fundamental requirement for practical system design. In [55], we showed that at high rates, under certain conditions, optimal quantizers are lattice quantizers, disconnected quantization cells need not be mapped into the same index, and there is asymptotically no performance loss by not having access to the side information at the encoder. This confirmed the experimental findings in [54].

### III. LOW-COMPLEXITY VIDEO ENCODING

Implementations of current video compression standards, such as the ISO MPEG schemes or the ITU-T recommendations H.263 and H.264 [56], [57] require much more computation for the encoder than for the decoder; typically the encoder is 5–10 times more complex than the decoder. This asymmetry is well suited for broadcasting or for streaming video-on-demand systems where video is compressed once and decoded many times. However, some applications may
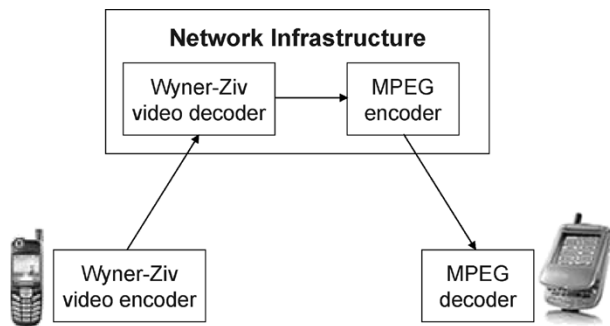
**Network Infrastructure**

Wyner-Ziv video decoder → MPEG encoder

Wyner-Ziv video encoder

MPEG decoder

**Fig. 6.** Transcoding architecture for wireless video.

require the dual system, i.e., low-complexity encoders, possibly at the expense of high-complexity decoders. Examples of such systems include wireless video sensors for surveillance, wireless PC cameras, mobile camera phones, disposable video cameras, and networked camcorders. In all of these cases, compression must be implemented at the camera where memory and computation are scarce.

The Wyner–Ziv theory [31], [37], [38], discussed in Section II-C, suggests that an unconventional video coding system, which encodes individual frames independently, but decodes them conditionally, is viable. In fact, such a system might achieve a performance that is closer to conventional interframe coding (e.g., MPEG) than to conventional intraframe coding (e.g., Motion-JPEG). In contrast to conventional hybrid predictive video coding where motion-compensated previous frames are used as side information, in the proposed system previous frames are used as "side information" at the decoder only.[1]

Such a Wyner–Ziv video coder would have a great cost advantage, since it compresses each video frame by itself, requiring only intraframe processing. The corresponding decoder in the fixed part of the network would exploit the statistical dependence between frames, by much more complex interframe processing. Beyond shifting the expensive motion estimation and compensation from the encoder to the decoder, the desired asymmetry is also consistent with the Slepian–Wolf and Wyner–Ziv coding algorithms, discussed in Sections II-B and II-D, which tend to have simple encoders, but much more demanding decoders.

Even if the receiver is another complexity-constrained device, as would be the case for video messaging or video telephony with mobile terminals at both ends, it is still advantageous to employ Wyner–Ziv coding in conjunction with a transcoding architecture depicted in Fig. 6. A mobile camera phone captures and compresses the video using Wyner–Ziv coding and transmits the data to the fixed part of the network. There, the bit stream is decoded and re-encoded using conventional video standards, such as MPEG. This architecture not only pushes the bulk of the computation into the fixed part of the network, but also shares the transcoder among many users, thus providing additional cost savings.

---

[1]Throughout this paper, the term "side information" is used in the spirit of distributed source coding. In the video coding literature, "side information" usually refers to transmitted motion vectors, mode decisions, etc. We avoid using the term in the latter sense to minimize confusion.

Both our own group at Stanford University, Stanford, CA [58]–[62], as well as Ramchandran's group at the University of California, Berkeley [63]–[65], have proposed practical schemes that are based on this novel video compression paradigm. We will review the coding algorithms and their performance in the following sections.

### A. Pixel-Domain Encoding

The simplest system that we have investigated is the combination of a pixel-domain intraframe encoder and interframe decoder system for video compression, as shown in Fig. 7 [58]–[60]. A subset of frames, regularly spaced in the sequence, serve as key frames $K$ which are encoded and decoded using a conventional intraframe $8 \times 8$ discrete cosine transform (DCT) codec. The frames between the key frames are "Wyner–Ziv frames" which are intraframe encoded but interframe decoded.

For a Wyner–Ziv frame $S$, each pixel value is uniformly quantized with $2^M$ intervals. We usually incorporate subtractive dithering to avoid contouring and improve the subjective quality of the reconstructed image. A sufficiently large block of quantizer indices $q$ is provided to the Slepian–Wolf encoder. The Slepian–Wolf coder is implemented using a rate-compatible punctured turbo code (RCPT) [66]. The RCPT provides the rate flexibility which is essential in adapting to the changing statistics between the side information and the frame to be encoded. In our system, the rate of the RCPT is chosen by the decoder and relayed to the encoder through feedback, as detailed below and in Section III-D.

For each Wyner–Ziv frame, the decoder generates the side information $\hat{S}$ by interpolation or extrapolation of previously decoded key frames and, possibly, previously decoded Wyner–Ziv frames. To exploit the side information, the decoder assumes a statistical model of the "correlation channel." Specifically, a Laplacian distribution of the difference between the individual pixel values $S$ and $\hat{S}$ is assumed. The decoder estimates the Laplacian parameter by observing the statistics from previously decoded frames.

The turbo decoder combines the side information $\hat{S}$ and the received parity bits to recover the symbol stream $q'$. If the decoder cannot reliably decode the original symbols, it requests additional parity bits from the encoder buffer through feedback. The "request-and-decode" process is repeated until an acceptable probability of symbol error is reached. By using the side information, the decoder often correctly predicts the quantization bin $q$ and it, thus, needs to request $k \leqslant M$ bits to establish which of the $2^M$ bins a pixel belongs to. Hence, compression is achieved.

After the receiver decodes the quantizer index $q'$, it calculates a minimum-mean-squared-error reconstruction $S' = E[S|q', \hat{S}]$ of the original frame $S$. If the side information $\hat{S}$ is within the reconstructed bin, the reconstructed pixel will take a value close to the side information value. However, if the side information and the decoded quantizer index $q'$ disagree, i.e., $\hat{S}$ is outside the quantization bin, the reconstruction function forces $\hat{S}$ to lie within the bin. It, therefore, limits the magnitude of the reconstruction error to a maximum value, determined by the quantizer coarseness. This
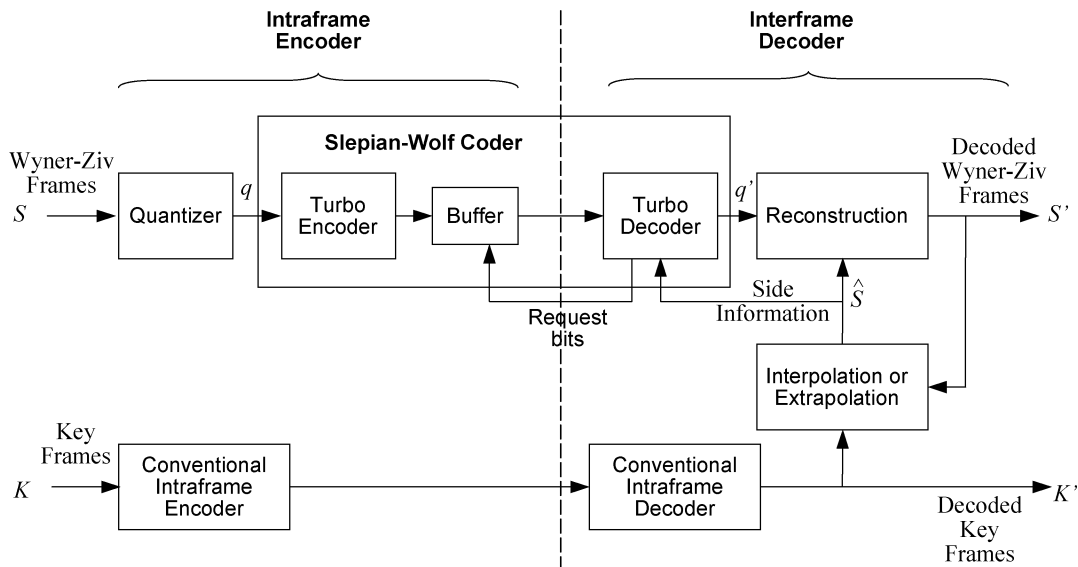
**Fig. 7.** Low-complexity video encoder and corresponding decoder.



**Fig. 8.** Sample frame from QCIF *Salesman* sequence. (a) Decoder side information generated by motion-compensated interpolation. (b) Reconstructed frame after Wyner–Ziv decoding. Note that motion-compensated interpolation gives good results for most of the image. The localized large interpolation errors are corrected by the Wyner–Ziv bit stream.

property is perceptually desirable, since it eliminates large errors which would be annoying to the viewer, as illustrated in Fig. 8. In this example, the side information, generated by motion-compensated interpolation, contains artifacts and blurring. Wyner–Ziv encoding [Fig. 8(b)] sharpens the image and reconstructs the hands and face even though the interpolation fails in these parts of the image.

Compared to motion-compensated predictive hybrid coding, pixel-domain Wyner–Ziv encoding is orders of magnitude less complex. Neither motion estimation and prediction, nor DCT and Inverse DCT (IDCT) are required at the encoder. The Slepian–Wolf encoder only requires two feedback shift registers and an interleaver. The interleaving is performed across the entire frame. In preliminary experiments on a Pentium III 1.2-GHz machine, we observe an average encoding runtime (without file operations) of

about 2.1 ms/frame for the Wyner–Ziv scheme, as compared to 36.0 ms/frame for H.263+ I-frame coding and 227.0 ms/frame for H.263+ B-frame coding.

Figs. 9 and 10 illustrate the rate-distortion performance of the pixel-domain Wyner–Ziv video coder for the sequences *Salesman* and *Hall Monitor*, in QCIF resolution with 10 frames/s. In these experiments, every fourth frame is a key frame and the rest of the frames are Wyner–Ziv frames. The side information is generated by motion-compensated interpolation using adjacent reconstructed frames. The plots show the rate and peak signal-to-noise ratio (PSNR), averaging over both key frames and Wyner–Ziv frames. We compare our results to: 1) DCT-based intraframe coding (all frames are encoded as I frames) and 2) H.263+ interframe coding with an I-B-B-B-I predictive structure. As can be seen from the plots, the pixel-domain Wyner–Ziv coder
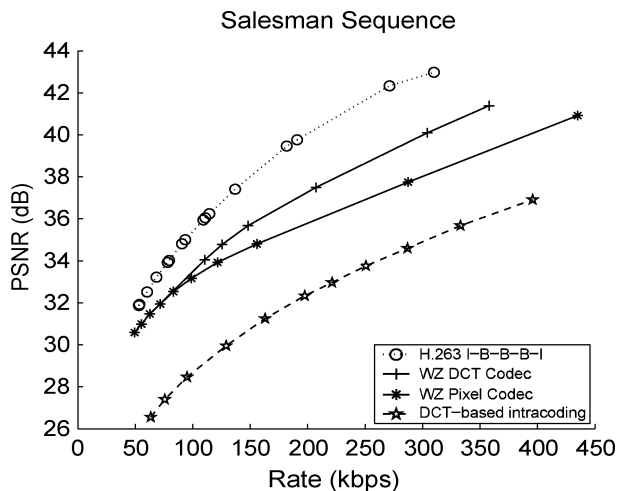
**Fig. 9.** Rate-distortion performance of the Wyner–Ziv video codec, compared to conventional intraframe video coding and interframe video coding. *Salesman* sequence.
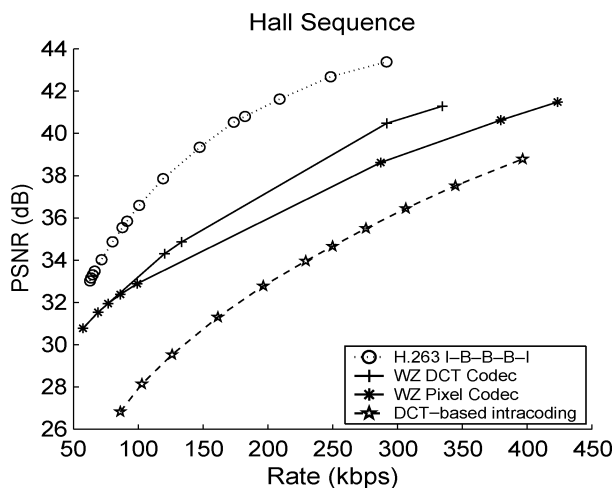


**Fig. 10.** Rate-distortion performance of the Wyner–Ziv video codec, compared to conventional intraframe video coding and interframe video coding. *Hall Monitor* sequence.

performs 2–5 dB better than conventional intraframe coding. There is still a significant gap toward H.263+ interframe coding.

### B. Transform-Domain Encoding

In conventional, nondistributed source coding, orthonormal transforms are widely used to decompose the source vector into spectral coefficients, which are individually coded with scalar quantizers and entropy coders. Pradhan and Ramchandran considered the use of transforms in Wyner–Ziv coding of images and analyzed the bit allocation problem for the Gaussian case [67]. Gastpar *et al.* [68], [69] investigated the Karhunen–Loève transform (KLT) for distributed source coding, but they assumed that the covariance matrix of the source vector given the side information does not depend on the values of the side information, and the study is not in the context of a practical coding scheme.

In our work [55], we theoretically studied the transformation of both the source vector and the side information,

with Wyner–Ziv coding of the subbands (Fig. 11) and established conditions under which the DCT is the optimal transform. We have implemented a Wyner–Ziv transform-domain video coder, closely following the block diagram (Fig. 11). First a block-wise DCT of each Wyner–Ziv frame is performed. The transform coefficients are independently quantized, grouped into coefficient bands, and then compressed by a Slepian–Wolf turbo coder [55], [61].

Similar to the pixel-domain scheme, the decoder utilizes previously reconstructed frames to generate a side information frame, with or without motion compensation. The correlation between a coefficient and the corresponding side information is also modeled using a Laplacian residual model with the parameters trained from different sequences. A blockwise DCT is applied to the side information frame, resulting in side information coefficient bands. A bank of turbo decoders reconstructs the quantized coefficient bands independently using the corresponding side information. Each coefficient band is then reconstructed as the best estimate given the reconstructed symbols and the side information.

In Figs. 9 and 10, we also plot the rate-distortion performance of a transform-domain Wyner–Ziv video codec using a 4 × 4 blockwise DCT. For both sequences, we observe a gain of up to 2 dB over the simpler pixel-based system. The spatial transform enables the codec to exploit the statistical dependencies within a frame, thus achieving better rate-distortion performance than the pixel-domain scheme.

The DCT-domain scheme has a higher encoder complexity than the pixel-domain system. It is similar to that of conventional intraframe video coding. However, it remains much less complex than interframe predictive schemes because motion estimation and compensation are not needed at the encoder.

A similar transform-domain Wyner–Ziv video coder has been developed independently by Puri and Ramchandran and presented under the acronym PRISM [63]–[65]. Like in our scheme, a blockwise DCT is followed by uniform scalar quantization. However, each block is then encoded independently. Only the low-frequency coefficients are compressed using a trellis-based Slepian–Wolf code. The higher-frequency coefficients are conventionally entropy-coded. The encoder also sends a cyclic redundancy check (CRC) of the quantized coefficients to aid motion compensation at the receiver. Preliminary rate-distortion results show a performance between conventional intraframe transform coding and conventional motion-compensated transform coding, similar to our own results.

### C. Joint Decoding and Motion Estimation

To achieve high compression efficiency in a Wyner–Ziv video codec, motion has to be estimated at the decoder. Conventional motion-compensated coding benefits from estimating the best motion vector by directly comparing the frame to be encoded with one or more reference frames. The analogous approach for Wyner–Ziv video coding requires joint decoding and motion estimation, using the Wyner–Ziv
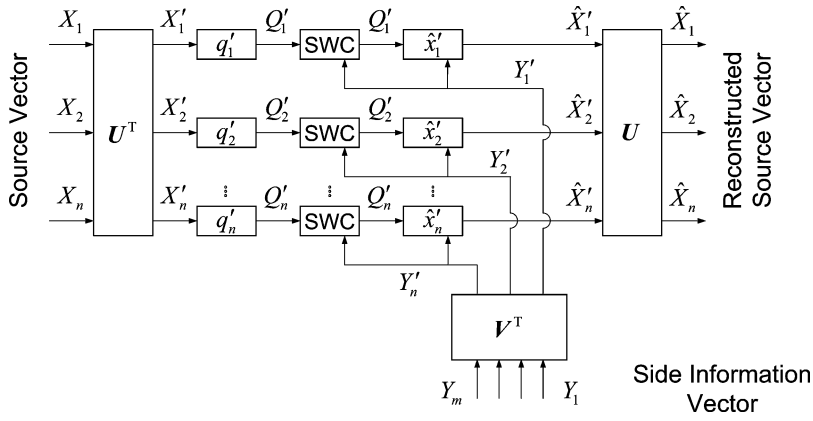
**Fig. 11.** Wyner–Ziv coding in the transform domain (SWC = Slepian–Wolf Coder).

bits, and possibly additional helper information from encoder. A first example of this has been implemented in the PRISM system, where the CRC of the quantized symbols aid in determining the motion at the decoder. Viterbi decoding is carried out for a set of motion-compensated candidate prediction blocks, each with a different motion vector. The CRC of each decoded version is then compared with the transmitted CRC to establish which version should be used [65].

In our own recent work, we send a robust hash code word to aid the decoder in estimating the motion [62]. Currently, our hash simply consists of a small subset of the quantized DCT coefficients. We apply this in a low-delay system where only the previous reconstructed frame is used to generate the side information of a current Wyner–Ziv frame. Since the hash is much smaller than the original data, we also allow the encoder to keep the hash code words for the previous frame in memory. For each block of the current frame, the distance to the corresponding robust hash of the previous frame is calculated. If the distance is small, a code word is sent that lets the decoder repeat this block from this frame memory. If the distance exceeds a threshold, the block's hash is sent, along with the Wyner–Ziv bits. Based on the hash, the decoder performs a motion search to generate the best side information block from the previous frame. The quantized coefficients of the hash code can fix the corresponding probabilities in the turbo decoder, thus further reducing the rate needed for the parity bits. The hash can also be utilized in refining the coefficients during reconstruction. Our approach is closely related to the idea of probing the correlation channel for universal Slepian–Wolf coding, proposed in [9].

The rate-distortion performance of the above low-delay system for different amounts of Wyner–Ziv frames between key frames [resulting in different lengths of group of pictures (GOP)] is shown in Figs. 12 and 13. Depending on the bit-rate, between about 5% and 20% of the hash codewords are sent. We compare the rate-distortion performance to that of H.263+ interframe coding (I-P-P) with the same GOP length. We substantially outperform conventional intraframe DCT coding, but there still remains a performance gap relative to H.263+ interframe coding.



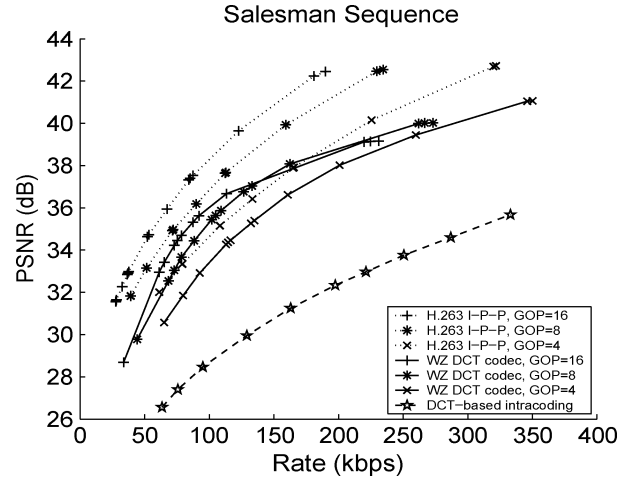**Fig. 12.** Rate-distortion performance of the Wyner–Ziv video codec with varying GOP length. *Salesman* sequence.
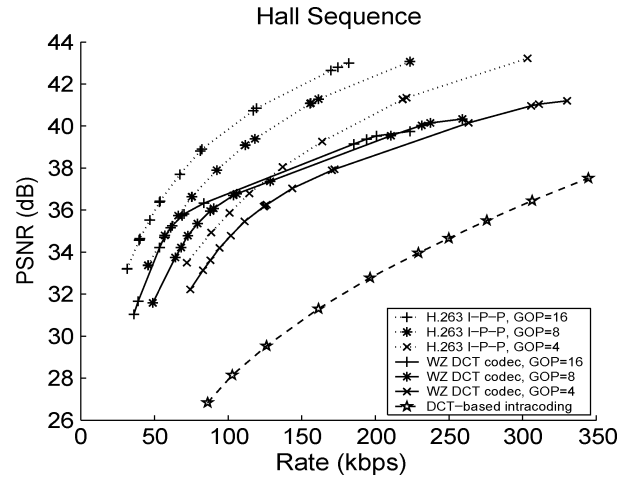


**Fig. 13.** Rate-distortion performance of the Wyner–Ziv video codec with varying GOP length. *Hall* sequence.

### D. Rate Control

The bit-rate for a Wyner–Ziv frame is determined by the statistical dependence between the frame and side information. While the encoding algorithm itself does not change,

the required bit-rate does as the correlation channel statistics change. The decision on how many bits to send for each frame is tricky, since the side information is exploited only at the decoder but not at the encoder.

One approach to solve the rate control problem relies entirely on the decoder and feedback information. The decoder determines the optimal encoding rate and sends this information to the encoder. In our pixel-domain and transform-domain schemes (Sections III-A and III-B), the decoder attempts decoding using the bits received so far. If the turbo decoding fails, the decoder requests additional bits from the encoder. As an alternative to this "decode-and-request" procedure, the decoder could implement a correlation channel estimation algorithm using previously reconstructed frames and send the predicted rate to the encoder.

Rate control performed by the decoder obviously minimizes the burden at the encoder. Feedback also allows the decoder to have great flexibility in generating the side information, from the very simple scheme of copying the previous frame to more intelligent algorithms such as motion-compensated interpolation with dense motion vector fields, object-based segmentation, or multiple frame predictors. If rate control is performed at the Slepian–Wolf decoder by requesting a sufficient number of bits, fewer bits are sent with more accurate side information. We can, therefore, improve the compression performance of the overall system by only altering the decoder. The encoder remains unchanged. This flexibility is the dual to conventional video compression, where decoders are fixed (usually by a standard), but the encoder has considerable flexibility to tradeoff smart processing versus bit-rate.

There are two obvious drawbacks of this rate control scheme. First, it requires a feedback channel and, thus, a higher latency is incurred. Second, the statistical estimation or the decode-and-request process has to be performed online, i.e., while the video is being encoded. Therefore, the algorithm is not suitable for applications such as low-complexity video acquisition-and-storage devices wherein the compressed video is transferred and decoded at a later time.

Another way to perform rate control is to allow some simple temporal dependence estimation at the encoder. For example, in the PRISM scheme [63]–[65], the encoder is permitted storage of one previous frame. Based on the frame difference energy, each block is classified into one of several coding modes. If the block difference is very small, the block is not encoded at all; if the block difference is large, the block is intracoded. In between these two extremes are various syndrome encoding modes with different rates, depending on the estimated statistical dependence. The rate estimation does not involve motion compensation and, hence, is necessarily inaccurate, if motion compensation is used at the decoder. Further, the flexibility of the decoder is restricted. Better decoder side information cannot lower the bit-rate, even though it can reduce reconstruction distortion. Most importantly, though, the rate control scheme of PRISM does not require online decoding and a feedback channel, thus making it suitable, for example, for storage applications.

## IV. ROBUST VIDEO TRANSMISSION

Wyner–Ziv coding can be thought of as a technique which generates parity information to correct the "errors" of the correlation channel between source sequence and side information, up to a distortion introduced by quantization. Wyner–Ziv coding, thus, lends itself naturally to robust video transmission as a "lossy" channel coding technique.

It is straightforward to use a stronger Slepian–Wolf code which not only corrects the discrepancies of the correlation channel, but additionally corrects errors introduced during transmission of the source sequence. Experiments have been reported for the PRISM codec [63]–[65] that compare the effect of frame loss with that observed in a conventional predictive video codec (H.263+). With H.263+, displeasing visual artifacts are observed due to interframe error propagation. With PRISM, the decoded video quality is minimally affected and there is no drift between encoder and decoder. Sehgal *et al.* [70] have also proposed a Wyner–Ziv coding scheme based on turbo codes to combat interframe error propagation. In their scheme, Wyner–Ziv coding is applied to certain "peg frames," while the remaining frames are encoded by a conventional predictive video encoder. This ensures that any decoding errors in the predictive video decoder can only propagate until the next peg frame. Jagmohan *et al.* [71] applied Wyner–Ziv coding to each frame to design a "state-free" video codec in which the encoder and decoder need not maintain precisely identical states while decoding the next frame. The state-free codec performs only 1–2.5 dB worse than a state-of-the-art standard video codec. Xu and Xiong used LDPC codes and nested Slepian–Wolf quantization to construct a layered Wyner–Ziv video codec [72]. This codec approaches the rate-distortion performance of a conventional codec with fine-granular scalability (FGS) coding and, more importantly, the LDPC code can be used to provide resilience to channel errors.

Wyner–Ziv coding is also the key ingredient of systematic lossy source-channel coding [73]. In this configuration, a source signal is transmitted over an analog channel without channel coding. An encoded version of the source signal is sent over a digital channel as enhancement information. The noisy received version of the original signal serves as side information to decode the output of the digital channel and produce the enhanced version as shown in Fig. 14. The term "systematic coding" has been introduced as an extension of systematic error-correcting channel codes to refer to a partially uncoded transmission. Shamai *et al.* established information-theoretic bounds and conditions for optimality of such a configuration in [73].

The application of systematic lossy source-channel coding to error-prone digital transmission is illustrated in Fig. 15. At the transmitter, the input video $S$ is compressed independently by an MPEG video coder and Wyner–Ziv encoder A. The MPEG video signal transmitted over the error-prone channel constitutes the systematic portion of the transmission, which is then augmented by the Wyner–Ziv bit stream. The Wyner–Ziv bit stream can be thought of as a second, independent description of $S$, but with coarser quantization.
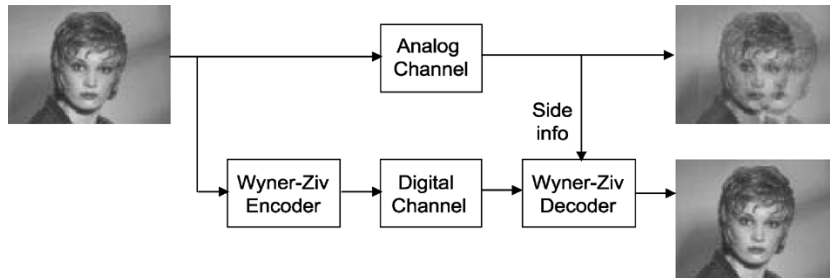
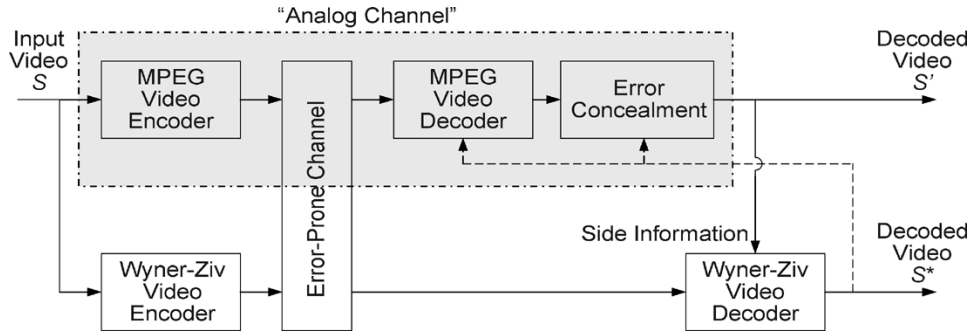**Fig. 14.** Digitally enhanced analog transmission.



**Fig. 15.** MPEG video transmission system in which the video waveform is protected by a Wyner–Ziv bit stream. Such a system achieves graceful degradation with increasing channel error rate without a layered signal representation. The systematic portion of the transmission corresponds to the analog channel of Fig. 14.

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 $S'$. We refer to the system in Fig. 15 as systematic lossy error protection (SLEP) in this paper.[2]

When transmission errors occur, the receiver performs error concealment, but some portions of $S'$ might still have unacceptably large errors. In this case, Wyner–Ziv bits allow reconstruction of the Wyner–Ziv description, using the decoded waveform $S'$ as side information. This coarser second description and side information $S'$ are combined to yield an improved decoded video $S^*$. In portions of $S'$ that are unaffected by transmission errors, $S^*$ is essentially identical to $S'$. However, in portions of $S'$ that are degraded by transmission errors, the coarser Wyner–Ziv representation limits the maximum degradation that can occur.

This system exploits the tradeoff between the Wyner–Ziv bit-rate and the residual distortion from transmission errors to achieve graceful degradation with worsening channel error rates. The conventional approach to achieve such graceful degradation employs layered video coding schemes [74]–[83] combined with unequal error protection, such as priority encoding transmission (PET) [84], [85]. Unfortunately, all layered video coding schemes standardized to date incur a substantial rate-distortion loss relative to single-layer schemes and, hence, are not used in practical systems. By not requiring a layered representation, the Wyner–Ziv scheme retains the efficiency of conventional nonscalable video codecs. The systematic scheme of Fig. 15 is com-

patible with systems already deployed, such as MPEG-2 digital TV broadcasting systems. The Wyner–Ziv bit streams can be ignored by legacy systems, but would be used by new receivers. As an aside, please note that other than in Section III, low encoder complexity is not a requirement for broadcasting, but decoder complexity might be a concern.

We presented first results of error-resilient video broadcasting with pixel-domain Wyner–Ziv coding [59], [86], and then with an improved Wyner–Ziv video codec which used a hybrid video codec with Reed–Solomon (RS) codes [87]. For a practical Wyner–Ziv codec, consider the forward error protection scheme shown in Fig. 16. The Wyner–Ziv codec 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 bit stream from this "coarse" encoder A, referred to as the Wyner–Ziv description, is input to a channel coder which applies RS codes across the slices of an entire frame. *Only the RS parity symbols* are transmitted to the receiver. When transmission errors occur, the conventionally decoded error-concealed video $S'$ is re-encoded by a coarse encoder B, which is a replica of the coarse encoder A. Therefore, the output of this coarse encoder B is identical to that of coarse encoder A except for those slices which are corrupted by channel errors. Since the locations of the erroneous slices are known, the RS decoder treats them as erasures and applies erasure decoding to recover the missing coarse slices. These coarse slices then replace the corrupted slices in the main video sequence. This 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 output $S^*$ of superior

---

[2]In previous publications, we have used the term forward error protection (FEP), but FEP can be easily confused with classic FEC.
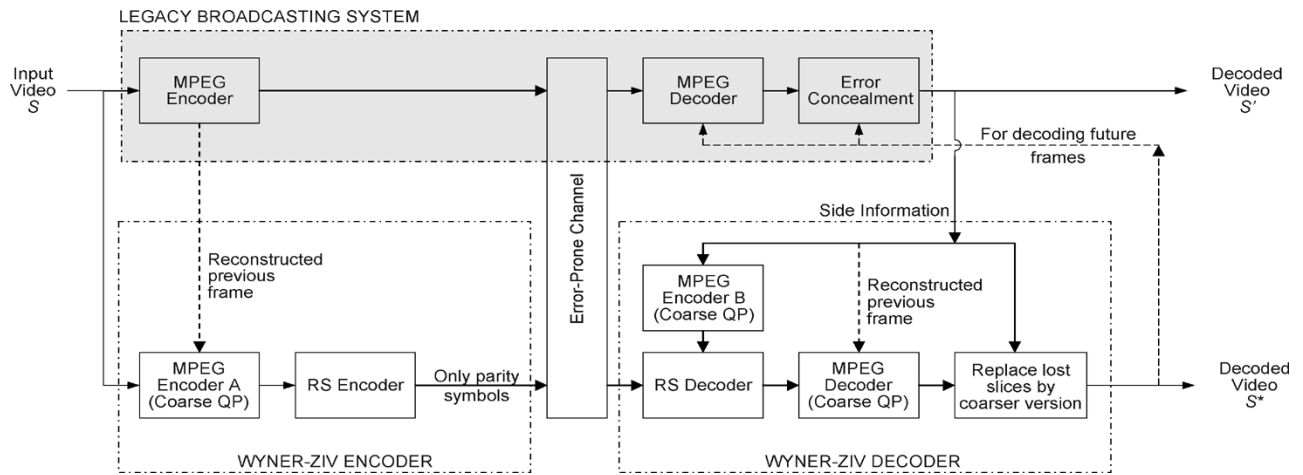
**Fig. 16.** Implementation of SLEP by combining hybrid video coding and RS codes across slices.
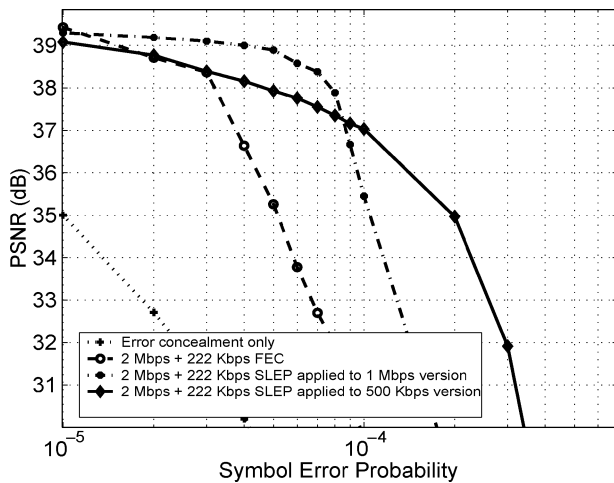


**Fig. 17.** SLEP uses Wyner–Ziv coding to trade off error correction capability with residual quantization mismatch, and provides graceful degradation of decoded video quality.

visual quality, using the conventionally decoded, error-concealed sequence as side information. To mitigate prediction mismatch between the two coarse video encoders A and B, the coarse encoder A must use the locally decoded previous frame from the "main" MPEG encoder, as a reference for predictive coding.

Fig. 17 shows a typical comparison between the SLEP scheme and traditional FEC. The systematic transmission consists of the Foreman CIF sequence encoded at 2 Mb/s with 222 kb/s of FEC applied to the systematic transmission, error correction breaks down at a symbol error probability of $3 \times 10^{-5}$. In the SLEP scheme, 222 kb/s of RS parity symbols are applied to a Wyner–Ziv description encoded at 1 Mb/s. This scheme registers a small drop in PSNR due to coarse quantization, but prevents further degradation of video quality until the error probability reaches $8 \times 10^{-5}$. Even higher error resilience is obtained if 222 kb/s of RS parity symbols are applied to an even coarser Wyner–Ziv description encoded at 500 kb/s. Visual examination of the decoded frames in Fig. 18 indicates that SLEP provides

significantly better picture quality than FEC at high error rates.

Note that the SLEP scheme described only requires the slice structure of the Wyner–Ziv video coder to match the typical transmission error patterns of the main video codec. Other than that, they can use entirely different compression schemes. For example, the main video codec could be legacy MPEG-2, while the Wyner–Ziv coder uses H.264 for superior rate-distortion performance. In practice, they would likely use the same compression scheme to avoid the coarse encoder B (Fig. 16) required in the Wyner–Ziv decoder, and use a much simpler transcoder in its place. We present such a simple transcoder in [88], which reuses the motion vectors and mode decisions from the main video codec and simply requantizes the DCT coefficients. The same transcoder is also used in the Wyner–Ziv encoder.

An interesting extension is the use of multiple embedded Wyner–Ziv streams in an SLEP scheme to shape the graceful degradation over a wider rate of channel error rates [88]. Such a system can choose to decode a finer or coarser Wyner–Ziv description depending on the channel error rate.

## V. CONCLUSION

Distributed coding is a fundamentally new paradigm for video compression. Based on Slepian and Wolf's and Wyner and Ziv's theorems from the 1970s, the development of practical coding schemes has commenced recently. Entropy coding, quantization, and transforms, the fundamental building blocks of conventional compression schemes, are now reasonably well understood in the context of distributed coding also. Slepian–Wolf encoding, the distributed variant of entropy coding, is related to channel coding, but fundamentally harder for practical applications due to the general statistics of the correlation channel. Further progress is needed toward adaptive and universal techniques. Design of optimal quantizers with receiver side information is possible with a suitable extension of the Lloyd algorithm. Quantizers with receiver side information may have non-connected quantizer cells, but in conjunction with optimum

(a)           (b)           (c)
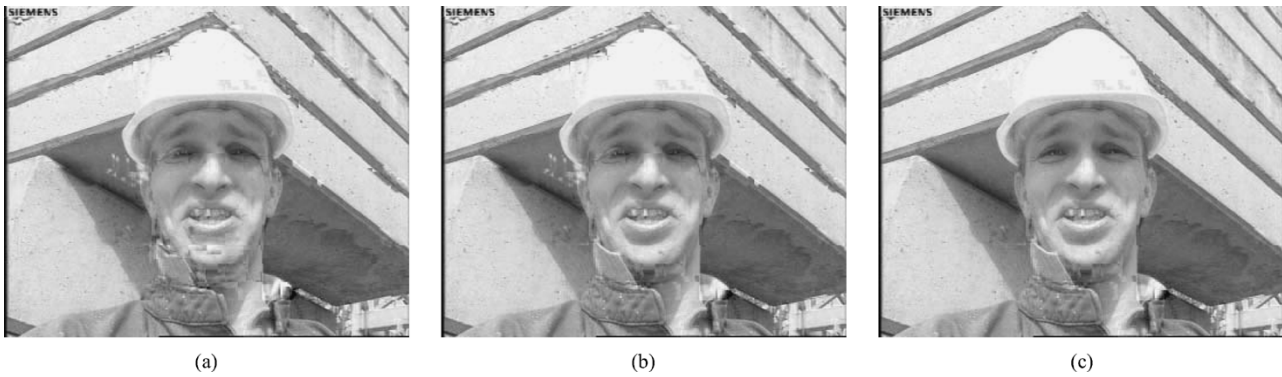
**Fig. 18.** Systematic lossy error protection provides substantially improved decoded video quality compared to FEC. A decoded frame from a channel trace at error probability $10^{-4}$ in Fig. 17 is shown here for visual comparison. (a) Error concealment only. (b) 222 kb/s FEC. (c) 222 kb/s SLEP for 1 Mb/s Wyner–Ziv description.

Slepian–Wolf coding, uniform quantizers often are appropriate. Transforms can be used with similar benefit as in conventional coding.

Wyner–Ziv coding, i.e., lossy compression with receiver side information, enables low-complexity video encoding, where the bulk of the computation is shifted to the decoder. Such schemes hold great promise for mobile video cameras, both for transmission and storage applications. We showed examples where the computation required for compression is reduced by 10X or even 100X relative to a conventional motion-compensated hybrid video encoder. With distributed coding, the interframe statistics of the video source are exploited at the decoder, but not at the encoder; hence, an intraframe encoder can be combined with interframe decoding. The rate-distortion performance of Wyner–Ziv coding does not yet reach the performance of a conventional interframe video coder, but it outperforms conventional intraframe coding by a substantial margin.

Its inherent robustness is a further attractive property of distributed coding, and joint source-channel coding is a natural application domain. Graceful degradation with deteriorating channel conditions can be achieved without a layered signal representation, thus overcoming a fundamental limitation of past schemes for video coding. The idea is to protect the original signal waveform by one or more Wyner–Ziv bit streams, rather than applying FEC to the compressed bit stream. Such a system can be built efficiently out of well-understood components, MPEG video coders and decoders and an RS coder and decoder.

It is unlikely that distributed video coding algorithm will ever beat conventional video coding schemes in rate-distortion performance, although they might come close. Why, then, should we care? Distributed video coding algorithms are radically different in regard to the requirements typically introduced by real-world applications, such as limited complexity, loss resilience, scalability, random access, etc. We believe that distributed coding techniques will soon complement conventional video coding to provide the best overall system performance and enable novel applications. From an academic perspective, distributed video coding offers a unique opportunity to revisit and reinvent most com-

pression techniques under the new paradigm, ranging from entropy coding to quantization, from transforms to motion compensation, from bit allocation to rate control. Much work remains to be done. If nothing else, this intellectual challenge will deepen our understanding of conventional video coding.

REFERENCES

[1] J. D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, pp. 471–480, Jul. 1973.
[2] A. D. Wyner, "Recent results in the Shannon theory," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 1, pp. 2–10, Jan. 1974.
[3] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): design and construction," in *Proc. IEEE Data Compression Conf.*, 1999, pp. 158–167.
[4] ——, "Distributed source coding: symmetric rates and applications to sensor networks," in *Proc. IEEE Data Compression Conf.*, 2000, pp. 363–372.
[5] ——, "Group-theoretic construction and analysis of generalized coset codes for symmetric/asymmetric distributed source coding," presented at the Conf. Information Sciences and Systems, Princeton, NJ, 2000.
[6] ——, "Geometric proof of rate-distortion function of Gaussian sources with side information at the decoder," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, 2000, p. 351.
[7] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 51–60, Mar. 2002.
[8] X. Wang and M. Orchard, "Design of trellis codes for source coding with side information at the decoder," in *Proc. IEEE Data Compression Conf.*, 2001, pp. 361–370.
[9] J. García-Frías, "Compression of correlated binary sources using turbo codes," *IEEE Commun. Lett.*, vol. 5, no. 10, pp. 417–419, Oct. 2001.
[10] J. García-Frías and Y. Zhao, "Data compression of unknown single and correlated binary sources using punctured turbo codes," presented at the Allerton Conf. Communication, Control, and Computing, Monticello, IL, 2001.
[11] J. Bajcsy and P. Mitran, "Coding for the Slepian–Wolf problem with turbo codes," in *Proc. IEEE Global Communications Conf.*, vol. 2, 2001, pp. 1400–1404.
[12] P. Mitran and J. Bajcsy, "Near Shannonlimit coding for the Slepian–Wolf problem," presented at the Biennial Symp. Communications, Kingston, ON, Canada, 2002.
[13] A. Aaron and B. Girod, "Compression with side information using turbo codes," in *Proc. IEEE Data Compression Conf.*, 2002, pp. 252–261.
[14] Y. Zhao and J. García-Frías, "Joint estimation and data compression of correlated nonbinary sources using punctured turbo codes," presented at the Conf. Information Sciences and Systems, Princeton, NJ, 2002.

[15] ——, "Data compression of correlated nonbinary sources using punctured turbo codes," in *Proc. IEEE Data Compression Conf.*, 2002, pp. 242–251.

[16] P. Mitran and J. Bajcsy, "Coding for the Wyner–Ziv problem with turbo-like codes," in *Proc. IEEE Int. Symp. Information Theory*, 2002, p. 91.

[17] ——, "Turbo source coding: a noise-robust approach to data compression," in *Proc. IEEE Data Compression Conf.*, 2002, p. 465.

[18] G. Zhu and F. Alajaji, "Turbo codes for nonuniform memoryless sources over noisy channels," *IEEE Commun. Lett.*, vol. 6, no. 2, pp. 64–66, Feb. 2002.

[19] J. García-Frías and Y. Zhao, "Compression of binary memoryless sources using punctured turbo codes," *IEEE Commun. Lett.*, vol. 6, no. 9, pp. 394–396, Sep. 2002.

[20] J. García-Frías, "Joint source-channel decoding of correlated sources over noisy channels," in *Proc. IEEE Data Compression Conf.*, 2001, pp. 283–292.

[21] A. Liveris, Z. Xiong, and C. Georghiades, "Joint source-channel coding of binary sources with side information at the decoder using IRA codes," presented at the Multimedia Signal Processing Workshop, St. Thomas, U.S. Virgin Islands, 2002.

[22] ——, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Commun. Lett.*, vol. 6, no. 10, pp. 440–442, Oct. 2002.

[23] ——, "Compression of binary sources with side information at the decoder using LDPC codes," presented at the IEEE Global Communications Symp., Taipei, Taiwan, R.O.C., 2002.

[24] D. Schonberg, S. S. Pradhan, and K. Ramchandran, "LDPC codes can approach the Slepian–Wolf bound for general binary sources," presented at the Allerton Conf. Communication, Control, and Computing, Champaign, IL, 2002.

[25] ——, "Distributed code constructions for the entire Slepian–Wolf rate region for arbitrarily correlated sources," in *Proc. Asilomar Conf. Signals and Systems*, 2003, pp. 835–839.

[26] ——, "Distributed code constructions for the entire Slepian–Wolf rate region for arbitrarily correlated sources," in *Proc. IEEE Data Compression Conf.*, 2004, pp. 292–301.

[27] T. P. Coleman, A. H. Lee, M. Medard, and M. Effros, "On some new approaches to practical Slepian–Wolf compression inspired by channel coding," in *Proc. IEEE Data Compression Conf.*, 2004, pp. 282–291.

[28] V. Stankovic, A. D. Liveris, Z. Xiong, and C. N. Georghiades, "Design of Slepian–Wolf codes by channel code partitioning," in *Proc. IEEE Data Compression Conf.*, 2004, pp. 302–311.

[29] C. F. Lan, A. D. Liveris, K. Narayanan, Z. Xiong, and C. Georghiades, "Slepian–Wolf coding of multiple M-ary sources using LDPC codes," in *Proc. IEEE Data Compression Conf.*, 2004, p. 549.

[30] N. Gehrig and P. L. Dragotti, "Symmetric and a-symmetric Slepian–Wolf codes with systematic and nonsystematic linear codes," *IEEE Commun. Lett.*, to be published.

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

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

[33] ——, "The rate-distortion function for source coding with side information at the decoder—II: general sources," *Inf. Control*, vol. 38, no. 1, pp. 60–80, Jul. 1978.

[34] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[35] J. K. Su, J. J. Eggers, and B. Girod, "Illustration of the duality between channel coding and rate distortion with side information," presented at the 2000 Asilomar Conf. Signals and Systems, Pacific Grove, CA.

[36] R. J. Barron, B. Chen, and G. W. Wornell, "The duality between information embedding and source coding with side information and some applications," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1159–1180, May 2003.

[37] S. Pradhan, J. Chou, and K. Ramchandran, "Duality between source coding and channel coding and its extension to the side information case," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1181–1203, May 2003.

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

[39] R. Zamir and S. Shamai, "Nested linear/lattice codes for Wyner–Ziv encoding," in *Proc. Information Theory Workshop*, 1998, pp. 92–93.

[40] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1250–1276, Jun. 2002.

[41] J. Kusuma, L. Doherty, and K. Ramchandran, "Distributed compression for sensor networks," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, vol. 1, 2001, pp. 82–85.

[42] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 51–60, Mar. 2002.

[43] S. D. Servetto, "Lattice quantization with side information," in *Proc. IEEE Data Compression Conf.*, 2000, pp. 510–519.

[44] Z. Xiong, A. Liveris, S. Cheng, and Z. Liu, "Nested quantization and Slepian–Wolf coding: a Wyner–Ziv coding paradigm for i.i.d. sources," presented at the IEEE Workshop Statistical Signal Processing (SSP), St. Louis, MO, 2003.

[45] Z. Liu, S. Cheng, A. Liveris, and Z. Xiong, "Slepian–Wolf coded nested quantization (SWC-NQ) for Wyner–Ziv coding: performance analysis and code design," in *Proc. IEEE Data Compression Conf.*, 2004, pp. 322–331.

[46] Y. Yang, S. Cheng, Z. Xiong, and W. Zhao, "Wyner–Ziv coding based on TCQ and LDPC codes," presented at the Asilomar Conf. Signals, Systems and Computers, Pacific Grove, CA, 2003.

[47] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Process. Mag.*, vol. 21, no. 5, pp. 80–94, Sep. 2004.

[48] M. Fleming, Q. Zhao, and M. Effros, "Network vector quantization," IEEE Trans. Inf. Theory, vol. 50, no. 8, pp. 1584–1604, Aug. 2004, submitted for publication.

[49] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. IT-28, pp. 129–1373, Mar. 1982.

[50] M. Fleming and M. Effros, "Network vector quantization," in *Proc. IEEE Data Compression Conf.*, 2001, pp. 13–22.

[51] D. Muresan and M. Effros, "Quantization as histogram segmentation: globally optimal scalar quantizer design in network systems," in *Proc. IEEE Data Compression Conf.*, 2002, pp. 302–311.

[52] M. Effros and D. Muresan, "Codecell contiguity in optimal fixed-rate and entropy-constrained network scalar quantizers," in *Proc. IEEE Data Compression Conf.*, 2002, pp. 312–321.

[53] J. Cardinal and G. V. Asche, "Joint entropy-constrained multiterminal quantization," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, 2002, p. 63.

[54] D. Rebollo-Monedero, R. Zhang, and B. Girod, "Design of optimal quantizers for distributed source coding," in *Proc. IEEE Data Compression Conf.*, Snowbird, UT, Mar. 2003, pp. 13–22.

[55] D. Rebollo-Monedero, A. Aaron, and B. Girod, "Transforms for high-rate distributed source coding," presented at the Asilomar Conf. Signals and Systems, Pacific Grove, CA, 2003.

[56] Y. Wang, J. Ostermann, and Y. Zhang, *Video Processing and Communications*, 1st ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[57] T. Wiegand, G. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[58] A. Aaron, R. Zhang, and B. Girod, "Wyner–Ziv coding of motion video," presented at the Asilomar Conf. Signals and Systems, Pacific Grove, CA, 2002.

[59] 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 Conf.*, 2003, pp. 93–102.

[60] A. Aaron, E. Setton, and B. Girod, "Toward practical Wyner–Ziv coding of video," presented at the IEEE Int. Conf. Image Processing, Barcelona, Spain, 2003.

[61] A. Aaron, S. Rane, E. Setton, and B. Girod, "Transform-domain Wyner–Ziv codec for video," presented at the SPIE Visual Communications and Image Processing Conf., San Jose, CA, 2004.

[62] A. Aaron, S. Rane, and B. Girod, "Wyner–Ziv video coding with hash-based motion compensation at the receiver," presented at the IEEE Int. Conf. Image Processing, Singapore, 2004.

[63] R. Puri and K. Ramchandran, "PRISM: a new robust video coding architecture based on distributed compression principles," presented at the Allerton Conf. Communication, Control, and Computing, Allerton, IL, 2002.

[64] ——, "PRISM: an uplink-friendly multimedia coding paradigm," presented at the Int. Conf. Acoustics, Speech, and Signal Processing, Hong Kong, China, 2003.

[65] ——, "PRISM: A 'reversed' multimedia coding paradigm," presented at the IEEE Int. Conf. Image Processing, Barcelona, Spain, 2003.

[66] D. Rowitch and L. Milstein, "On the performance of hybrid FEC/ARQ systems using rate compatible punctured turbo codes," *IEEE Trans. Commun.*, vol. 48, no. 6, pp. 948–959, Jun. 2000.

[67] 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 Conf.*, 2001, pp. 63–72.

[68] M. Gastpar, P. L. Dragotti, and M. Vetterli, "The distributed Karhunen–Loève transform," in *Proc. IEEE Int. Workshop Multimedia Signal Processing (MMSP)*, 2002, pp. 57–60.

[69] M. Gastpar, P. Dragotti, and M. Vetterli, "The distributed, partial, and conditional Karhunen–Loève transforms," in *Proc. IEEE Data Compression Conf.*, 2003, pp. 283–292.

[70] A. Sehgal and N. Ahuja, "Robust predictive coding and the Wyner–Ziv problem," in *Proc. IEEE Data Compression Conf.*, 2003, pp. 103–112.

[71] A. Sehgal, A. Jagmohan, and N. Ahuja, "A causal state-free video encoding paradigm," presented at the IEEE Int. Conf. Image Processing, Barcelona, Spain, 2003.

[72] Q. Xu and Z. Xiong, "Layered Wyner–Ziv video coding," *IEEE Trans. Image Process.*, submitted for publication.

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

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

[75] A. Mohr, E. Riskin, and R. Ladner, "Unequal loss protection: graceful degradation of image quality over packet erasure channels through forward error correction," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 819–828, Jun. 2000.

[76] U. Horn, K. Stuhlmüller, M. Link, and B. Girod, "Robust internet video transmission based on scalable coding and unequal error protection," *Image Commun. (Special Issue on Real-Time Video Over the Internet)*, vol. 15, no. 1–2, pp. 77–94, Sep. 1999.

[77] Y.-C. Su, C.-S. Yang, and C.-W. Lee, "Optimal FEC assignment for scalable video transmission over burst-error channels with loss rate uplink," presented at the Packet Video Workshop, Nantes, France, Apr. 2003.

[78] W. Tan and A. Zakhor, "Error control for video multicast using heirarchical FEC," in *Proc. Int. Conf. Image Processing*, 1999, pp. 401–405.

[79] W. Heinzelman, M. Budagavi, and R. Talluri, "Unequal error protection of MPEG-4 compressed video," in *Proc. Int. Conf. Image Processing*, vol. 2, 1999, pp. 503–504.

[80] Y. Wang and M. D. Srinath, "Error resilient packet video with unequal error protection," presented at the Data Compression Conf., Snowbird, UT, 2001.

[81] G. Wang, Q. Zhang, W. Zhu, and Y.-Q. Zhang, "Channel adaptive unequal error protection for scalable video transmission over wireless channel," presented at the Visual Communications and Image Processing (VCIP 2001), San Jose, CA.

[82] W. Xu and S. S. Hemami, "Spectrally efficient partitioning of MPEG video streams for robust transmission over multiple channels," presented at the Int. Workshop Packet Video, Pittsburgh, PA, 2002.

[83] M. Andronico, A. Lombardo, S. Palazzo, and G. Schembra, "Performance analysis of priority encoding transmission of MPEG video streams," in *Proc. Global Telecommunications Conf. (GLOBECOM '96)*, 1996, pp. 267–271.

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

[85] S. Boucheron and M. R. Salamatian, "About priority encoding transmission," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 699–705, Mar. 2000.

[86] A. Aaron, S. Rane, D. Rebollo-Monedero, and B. Girod, "Systematic lossy forward error protection for video waveforms," in *Proc. IEEE Int. Conf. Image Processing*, 2003, pp. I-609–I-612.

[87] S. Rane, A. Aaron, and B. Girod, "Systematic lossy forward error protection for error resilient digital video broadcasting," presented at the SPIE Visual Communications and Image Processing Conf., San Jose, CA, 2004.

[88] ——, "Systematic lossy forward error protection for error resilient digital video broadcasting—a Wyner–Ziv coding approach," presented at the IEEE Int. Conf. Image Processing, Singapore, Oct. 2004.

**Bernd Girod** (Fellow, IEEE) received the M.S. degree from Georgia Institute of Technology (Georgia Tech), Atlanta, in 1980 and the Engineering Doctorate from University of Hannover, Hannover, Germany, in 1987.

He was Chaired Professor of Telecommunications in the Electrical Engineering Department of the University of Erlangen-Nuremberg, Germany, from 1993 to 1999. Prior visiting or regular faculty positions include the Massachusetts Institute of Technology, Cambridge; Georgia Tech; and Stanford University, Stanford, CA. He is currently Professor of Electrical Engineering and (by courtesy) Computer Science in the Information Systems Laboratory, Stanford University. He has published over 300 scientific papers, three monographs, and a textbook. He has also worked with several startup ventures as Founder, Director, Investor, or Advisor, among them Vivo Software, $8 \times 8$ (Nasdaq: EGHT), and RealNetworks (Nasdaq: RNWK). His research interests are in the areas of video compression and networked media systems.

Prof. Girod is a recipient of the 2004 EURASIP Technical Achievement Award.

**Anne Margot Aaron** received B.S. degrees in physics and in computer engineering from Ateneo de Manila University, Manila, The Philippines, in 1998 and 1999, respectively, and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 2001. She is currently working toward the Ph.D. degree at Stanford University.

Her research interests include distributed source coding, video compression, and multimedia systems.

**Shantanu Rane** (Student Member, IEEE) received the B.E. degree in instrumentation engineering from the Government College of Engineering, Pune University, Pune, India, in 1999 and the M.S. degree in electrical engineering from the University of Minnesota, Minneapolis, in 2001. He is currently working toward the Ph.D. degree in electrical engineering at Stanford University, Stanford, CA.

His research interests include distributed video coding and image communication in error-prone environments.

**David Rebollo-Monedero** received the B.S. degree in telecommunications engineering from the Technical University of Catalonia (UPC), Barcelona, Spain, in 1997 and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 2003. He is currently working toward the Ph.D. degree in electrical engineering at Stanford University.

He was with PricewaterhouseCoopers, Barcelona, Spain, as an Information Technology Consultant from 1997 to 2000. His research interests include quantization, transforms, and statistical inference for distributed source coding with video applications.