

Combined Forward Error Control and Packetized Zerotree Wavelet Encoding for Transmission of Images Over Varying Channels

Pamela C. Cosman, *Member, IEEE*, Jon K. Rogers, P. Greg Sherwood, *Member, IEEE*, and Kenneth Zeger, *Fellow, IEEE*

Abstract—One method of transmitting wavelet based zerotree encoded images over noisy channels is to add channel coding without altering the source coder. A second method is to reorder the embedded zerotree bitstream into packets containing a small set of wavelet coefficient trees. We consider a hybrid mixture of these two approaches and demonstrate situations in which the hybrid image coder can outperform either of the two building block methods, namely on channels that can suffer packet losses as well as statistically varying bit errors.

Index Terms—Image coding, joint source channel coding, packet networks, quality evaluation, wavelet zerotrees.

NOMENCLATURE

AWGN	Additive white Gaussian noise. An analog memoryless channel.
BER	Bit error rate. Average bit error probability of the channel.
BPSK	Binary phase-shift keying. A two symbol modulation scheme.
BSC	Binary symmetric channel. A discrete memoryless channel with a single parameter determining the bit error probability.
CRC	Cyclic redundancy check. A shortened cyclic code typically used for error detection.
EZW	Embedded zerotree wavelet. Shapiro's embedded wavelet-based image coder [4].
FEC	Forward error control.
MSE	Mean squared error. A fidelity criterion.
PSNR	Peak signal-to-noise ratio. A common fidelity measure with a logarithmic scale used in image coding.
PZW	Packetized zerotree wavelet. Robust zerotree coder designed for packet erasure channels by Rogers and Cosman [11].
RCPC	Rate-compatible punctured convolutional. A class of convolutional error correcting codes which allow

SNR	Signal-to-noise ratio. A measure of the relative strength of the signal versus the noise.
SPIHT	Set partitioning in hierarchical trees. An embedded wavelet-based image coder by Said and Pearlman [3].
UEP	Unequal error protection. Refers to applying different levels of channel coding to the source bits based on their importance to the decoding.

I. INTRODUCTION

THE transmission of images across noisy channels is fundamentally important in many applications and is still an unsolved problem for many types of channels. There has been some progress on this problem recently for certain specific channel conditions. For example, in [1] and [2], a concatenated channel coding scheme was applied to the set partitioning in hierarchical trees (SPIHT) [3] image coding algorithm (an improved version of the embedded zerotree wavelet (EZW) algorithm [4]) to achieve substantial performance gains over previous image coding systems, as long as the images are transmitted across memoryless channels [e.g., a binary symmetric channel (BSC)] with known statistics. The technique in [1] and [2] was used in [5] to provide error protection for video coding. In [6], an effective coding procedure that builds upon the method in [1], [2] was developed for transmission channels that allow feedback. Channels with and without feedback were considered in [7], where error detection is accomplished by introducing redundancy in an arithmetic source coder instead of through channel codes. In [8], a class of modified EZW algorithms was presented which limit error propagation by reducing the amount of variable length coding in the transmitted bitstream. A video codec proposed in [9] uses entropy coded scalar quantization of subband coefficient trees together with run-length and Huffman coding to produce a robust source coder. The bitstream in [9] is divided into variable-length independent packets which allow complete subband coefficient trees to either be lost or received as a unit. In [10] and [11], the output bitstream produced by the zerotree encoders of [3] and [4] was reordered and packetized in such a way that complete trees of wavelet coefficients were contained within packets. This allows graceful degradation of an image in the presence of packet erasures, instead of loss of synchronization

Manuscript received July 21, 1998; revised November 16, 1999. This work was supported in part by the National Science Foundation and by the Center for Wireless Communications, UCSD. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yoshitaka Hashimoto.

The authors are with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407 USA (e-mail: {pcosman, jkrogers, sherwood, zeger}@code.ucsd.edu; URL: <http://code.ucsd.edu>).

Publisher Item Identifier S 1057-7149(00)04869-7.

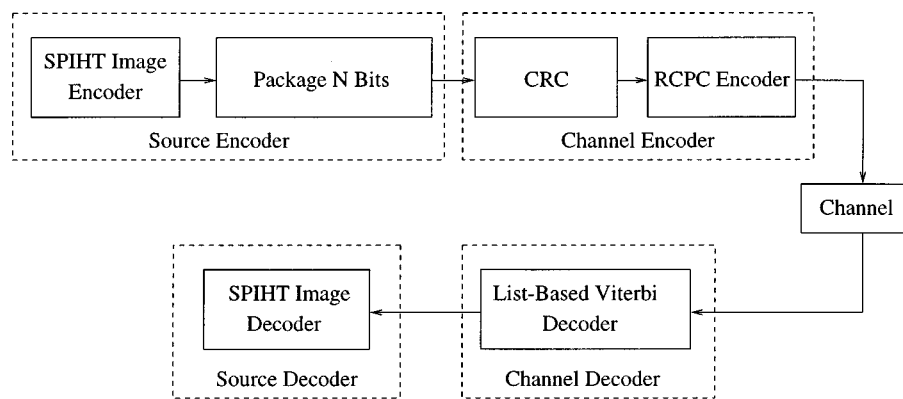


Fig. 1. Block diagram of the SPIHT+RCPC/CRC image coder.

typically experienced with the error-sensitive zerotree encoder. It is also somewhat robust to the effects of bursty channels. A similar method for encoding wavelet coefficient trees in groups was proposed in [12], where groups were independently compressed and interleaved for transmission. Because groups are decoded separately, error propagation is limited to a single group. In [13], a product channel code structure is used to make the system in [1] and [2] robust to fading channels and also slightly better on memoryless channels. In [14], source redundancy is used to improve the performance of an image coder transmitting over a Gilbert–Elliot channel. In [15], a multistage encoding structure is used to provide robustness to packet losses and bit errors.

Many data transmission environments are characterized by unknown and highly varying channel conditions. The mobile wireless environment is one example where channel conditions vary widely in a time span dependent on the mobile velocity. Since the mobile speed and its surrounding terrain may vary during transmission, it is difficult to accurately measure channel conditions and to adapt the coding. Also, in certain situations, such as a broadcast channel, it is impossible to perfectly match the channel since each receiver experiences a different channel. In these situations, it is important for the coding method to maintain adequate performance across a range of operating conditions and to degrade gracefully as conditions worsen.

For certain systems, another source of transmission impairment is packet loss due to such things as buffer overflow, misrouting, or unacceptably long arrival delays. In situations where wireless and wireline networks are connected, packets can be dropped due to queue overflow at the interface to a shared resource such as a base station transmitter. These situations can be modeled by the packet erasure channel. In reality, a mobile receiver may experience both packet losses and also bit errors on those packets which are not lost. It is precisely this combination of channel impairments that we address in this paper by introducing a robust hybrid encoding scheme.

While systems using forward error control (FEC) provide good protection for known channel conditions, if a precise statistical description of the channel is unknown, then one typically designs a FEC code for the worst possible channel that can be anticipated. Similarly, for existing robust image coding systems, one generally pays a significant source coding penalty when the

channel is clear. The fact that these two extreme coding schemes have such drawbacks motivates the present work.

We present the idea of combining FEC with a packetized error-resilient source coding scheme, in order to achieve a more robust image coding system. The goal is to introduce a hybrid coding approach that can survive a very poor channel, both in terms of high bit error rates and packet erasure rates, while providing good performance when channel noise is not significant.

We demonstrate the hybrid approach by describing a combination of two existing systems for transmitting images over noisy channels. In particular, we combine the FEC method of [1] and [2] with the zerotree wavelet packetization method of [10] and [11] in a hybrid structure that provides more robust performance over varying channel conditions, than either of the two methods by themselves. We measure the performance improvement of the hybrid coder on a channel which suffers bit errors as well as packet erasure. The particular hybrid coder presented is not claimed to be optimal, but rather was chosen as an example of the potential improvement possible using this new design approach. Combining improved source coding and channel coding systems will likely lead to better performance, although complexity may increase.

We describe in Section II two current methods of image coding for noisy channels. Sections III and IV explain how these current methods are combined to produce the hybrid coder. Section V describes the implementation of the channel model used for the performance tests. Results and conclusions from those tests are presented in Section VI. Code parameters are provided in the Appendix.

II. INDIVIDUAL COMPONENTS

A. Separate Source and Channel Coding

The robust image coding method presented in [1] and [2] used an efficient source coding algorithm (SPIHT [3]) followed by a strong concatenated channel code (RCPC/CRC). A block diagram of the system is shown in Fig. 1. This system is an example of an image coder based on Shannon’s “separation principle,” where the source coder is designed to maximize compression performance without regard for the channel, while the channel coder is designed to minimize error probability uniformly for the decoded bits. The combination proved to be very effective

in [1] and [2] for transmission over a binary symmetric channel under known channel conditions.

The channel code in [1] and [2] is a concatenation of an outer cyclic redundancy code (CRC) and an inner rate-compatible punctured convolutional (RCPC) code [16]. The code structure allows flexibility in selecting parameters such as block length (packet size), code rate, and error correction capability/complexity. The outer CRC serves a dual purpose by providing improved error correction performance when used in the list-Viterbi decoding of the RCPC code, as well as a high probability indication of channel decoding failure (i.e., error detection) which the source decoder can use to minimize the effects of uncorrected errors.

The excellent compression performance of the SPIHT source coder comes at the expense of a significant sensitivity to channel errors. Bit errors often lead to a complete loss of synchronization in the decoder due to the use of variable length coding and significant amounts of state information. The decoder essentially needs to decode the bits in a sequential and uninterrupted fashion for correct interpretation of the bits. However, the performance of the SPIHT+RCPC/CRC coder is good for the BSC with a known error rate because the FEC can efficiently lower the probability of decoding error. Also, the CRC allows detection of uncorrected packets so the source decoder can stop decoding before errors propagate and corrupt the image. This strategy often results in acceptable image quality because of the progressive nature of the source coder. For this type of channel, existing robust source coders are not as efficient despite the fact that additional source rate is available due to the need for less channel coding.

Difficulties occur on variable and unknown channels or on erasure channels. Since the source decoder in [1] and [2] stops decoding at the first uncorrectable error, the FEC must provide a low probability of decoding error, in order to achieve good performance. Channel codes typically transition rapidly from the designed performance to the uncoded performance (and even worse) as the channel degrades [17]. This lack of graceful degradation means that the FEC may need to be designed for a worst case channel, which thus sacrifices performance when the channel is clear (i.e., too little of the rate is spent on source coding under good channel conditions, especially for highly variable channels).

B. Robust Source Coding

Source coding can be designed to provide noise robustness without explicit error-correction coding. The packetized zerotree wavelet (PZW) coder [10] provides robustness by producing a compressed image datastream consisting of independently decodable packets.

PZW is an error-resilient variation on the EZW and SPIHT coders [3], [4]. Using bitplane encoding of the wavelet coefficients, the encoder generates substreams corresponding to individual coefficient trees (using the same tree structure as [4]). Groups of substreams are placed together into fixed-length packets (e.g., we use 384 bits). To facilitate this grouping into short packets, PZW uses no arithmetic coding and has fewer levels of wavelet decomposition than SPIHT. Because individual tree rates vary, the number of trees per packet varies,

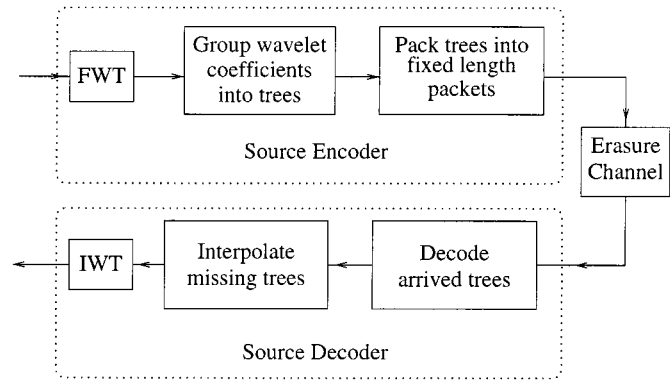


Fig. 2. Block diagram of the PZW image coder.

and the trees are either grown or pruned as necessary to fill each packet exactly. A header in each packet identifies the first tree and the number of trees in the packet. Within each packet, bitplanes of the trees are interleaved as in SPIHT, allowing the decoder to interpret the data with no additional overhead about the sizes and rates of individual trees. Fig. 2 shows a block diagram of the PZW algorithm. Coefficient trees in a correctly received packet are decodable independent of any other packet. Missing trees are concealed in the low-low wavelet band by interpolation from their immediate neighbors, and missing coefficients in higher bands are set to zero prior to inverse wavelet transforming the array.

The PZW scheme was designed with packet erasures in mind. Since each packet is independently decodable, lost packets do not lead to loss of synchronization between the encoder and decoder. For channels which contain bit errors as well as packet erasures, a 16 bit CRC can be added to each packet. Packets with detected errors are discarded by the source decoder. This error detection makes PZW robust over a bursty bit error channel. But transmission over a channel with more uniform errors (like a BSC) will result in more packet discards than over a channel with bursty errors. For example, assume we want to transmit an image which has been compressed to 50 000 bits and grouped into packets of length 400 (a total of 125 packets). Assume that there are 400 bit errors during transmission. For a BSC, 400 bit errors ($BER = 0.008$) translates to one error in each 125 bits on average, so each packet receives more than 3 bit errors (on average) and the entire stream is lost. However, a bursty channel might produce that same number of errors in 1000 bits, leading to only three or four lost packets.

The packetizing operations lead to a source coding performance loss of about 0.5–0.7 dB in peak signal-to-noise ratio (PSNR) relative to the SPIHT coder (without arithmetic coding) in an error-free case, but they provide robustness in the presence of channel noise. Errors cannot propagate beyond packet boundaries. Packets are of equal importance; given a certain packet loss rate, it matters little to the final PSNR which packets were lost.

III. HYBRID CODER

The hybrid coder consists of the PZW source coder (i.e., the source coder in Fig. 2) combined with the RCPC/CRC FEC

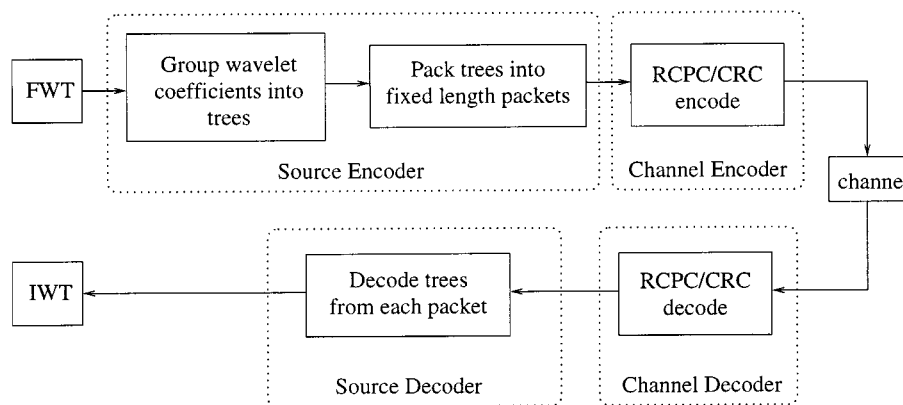


Fig. 3. Block diagram of the proposed coder.

coder (i.e., the channel coder in Fig. 1). Fig. 3 shows a block diagram of the proposed method. Image data is transformed into the wavelet domain. Groups of coefficient trees are placed into fixed length packets (as in PZW). Each packet is then wrapped by a RCPC/CRC code (as in the SPIHT+RCPC/CRC algorithm). The protected packets are sent across the channel and decoded by the channel decoder followed by the source decoder. Missing wavelet coefficients in the low-low band (from packet erasures or channel decoding failures) are replaced by interpolating values of all available 8-neighbor pixels. Missing coefficients in higher bands are set to zero before inverse wavelet transforming.

The two approaches used together are intended to help exploit the advantages of each other. Bit errors occurring throughout the stream (which would devastate the PZW coder) are corrected by the FEC; severe bursts of errors (which would overwhelm the FEC or would impose too severe a rate penalty if included in the FEC design) are absorbed by the underlying resilience of the PZW source coder. The PZW coder also handles packet erasures which would cause early truncation of the bitstream for the SPIHT coder.

The proposed coder is designed for a channel (discussed in detail in Section V) consisting of a wireline portion and a wireless portion. Losses in the wireline portion consist of packet erasures where lost packets, due to buffer overflow or mis-routing, do not arrive at the receiver. The wireless portion of the channel causes losses from excessive bit errors due to the fading channel characteristic.

In particular, the packet length at the output of the source encoder (used on the wireline portion of the channel) is fixed at 384 bits to match ATM packet sizes. The RCPC/CRC code is not designed for the worst channel, but rather for conditions in the middle of the expected range for the wireless portion of the channel. Thus, more transmission rate can be dedicated to source coding than with the SPIHT+RCPC/CRC coder which would have to be designed for the worst case. The method for selecting the appropriate channel code is described in Section IV.

Although the proposed method is not claimed to be optimal, the particular components are well suited for each other and for the goal of producing a robust coder on the combined packet erasure/bit error channel. First, the PZW coder works well for the small ATM packet lengths (i.e., the source coding perfor-

mance is close to that without segmentation into packets). The small packet size helps provide robustness since a packet only contains a small portion of the image information which means the impact of a lost packet is small. Second, the RCPC/CRC FEC provides good error correction performance for the wireless portion of the channel which introduces bit errors. As part of the decoding process, channel decoding failures can be detected with high probability so the source decoder can drop uncorrected packets rather than corrupt the image by decoding those bits. Use of more recent FEC methods such as turbo codes provides the most benefits in situations where soft decision channel decoding is available, channel state information is fairly well known, and the added decoding complexity is acceptable. The improved performance of these codes typically requires longer block lengths which may reduce robustness when channel conditions are worse than the design conditions. Also it is important to detect channel decoding failures, so additional coding would be necessary if this capability is not normally a feature of the code.

IV. PARAMETER VALUE OPTIMIZATION

In evaluating the performance of an image coder for noiseless channels, the PSNR versus bit rate curve is often used. In general, coder A is said to perform better than coder B if the curve of PSNR versus bit rate for A is pointwise greater than that for B. If the comparison between coders is performed at a specific bit rate, then the results can always be ordered, and the system designer has a clear optimization goal. But a single transmission rate does not capture the overall behavior of a coder used progressively, and may thus be misleading. Without focusing on a particular bit rate, there may not be a clear winner, since the curves may cross. Also, the optimization goal is difficult to specify when considering entire curves, since the performance at low bit rates or at high bit rates may be more important for specific applications.

In the noisy channel case, the situation is more complicated. Suppose the channel code rate and total bit rate are fixed. Then, the performance of the system is characterized by the cumulative distribution function (CDF) of the mean squared error (MSE), denoted by $F(x) = \text{Prob}(\text{Decoded Distortion} < x)$. The goal is to design a system with a high probability of pro-

ducing images with low MSE. Thus, a CDF is good if it rises sharply at a low MSE. As before, we would say that system A outperforms system B if the CDF for A lies everywhere above the CDF for B. If the two CDF curves cross, the comparative evaluation is more difficult. As before, one can try to escape this complication by looking at a particular point, say, the median MSE. But this method is unsatisfactory, since it ignores overall behavior.

In this section, we describe a general class of performance measures which can be specialized for the specific design goals of an application. Given a performance measure and a range of channel conditions, the best choice of channel code for both the SPIHT+RCPC/CRC coder and the hybrid coder can be found. Let Φ be a specification of the system free parameters to be optimized. For our case, Φ is a channel code (or equivalently the corresponding channel code rate, when a family of codes parameterized by code rate is used). For any particular channel code, the performance of the system is characterized by the cumulative distribution function (CDF) of the mean squared error (MSE), denoted by $F_{\Phi}(x) = \text{Prob}(\text{Decoded Distortion} < x)$. In order to quantify the performance of a system, it is convenient to reduce the function $F_{\Phi}(x)$ to a single real number. We thus define the fidelity criterion

$$J_{\Phi} = \int_0^{\infty} W(x)(1 - F_{\Phi}(x)) dx \quad (1)$$

where $W(x)$ is a nonnegative “weighting” function which allows us to emphasize or deemphasize the contribution of a particular MSE region to the total fidelity. In general, it results in a fidelity criterion which is a weighted conditional mean of the decoded MSE values. The goal is to minimize J_{Φ} over all Φ in some predefined family of codes.

Some specific examples are given below, where the unit step function is $u(x) = 1$ for $x \geq 0$ and $u(x) = 0$ for $x < 0$, and where $\delta(x)$ is the Dirac delta-function.

- *Example (i):* $W(x) = 1$

This fidelity criterion is the mean decoded MSE. Fig. 4 demonstrates this example, where J_{Φ} is the shaded area in the graph.

- *Example (ii):* $W(x) = u(-x + x_h) + C\delta(x - x_h)$

This fidelity criterion assumes a maximum level x_h of tolerable MSE for a decoded image. Images decoded with distortion greater than x_h are useless and therefore should not influence the optimization with respect to images having smaller MSE. The constant C allows the designer to trade off between the weighted conditional expected distortion of images decoded below x_h and the probability that images will be decoded above x_h . Fig. 5 shows an example with such a weighting function where the fidelity criterion is equal to the area of the shaded region.

- *Example (iii):* $W(x) = u(x - x_l)$

This fidelity criterion assumes any MSE below a certain threshold x_l is essentially equal to zero. This can be due to limitations of the output device or limits of human perception, for example. Fig. 6 shows an example where this fidelity criterion is equal to the area of the shaded region.

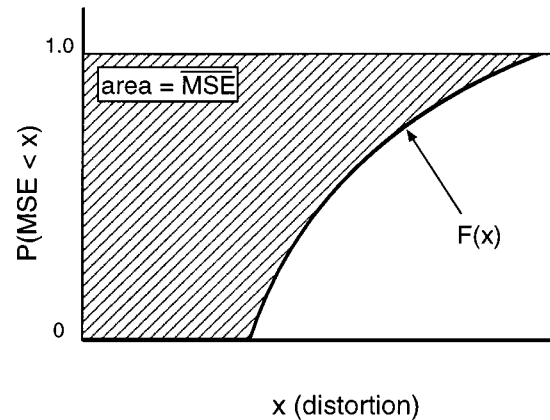


Fig. 4. Shaded area equals the fidelity criterion J_{Φ} with $W(x) = 1$ which is the mean decoded MSE in this case. $F(x)$ in the graph represents a generic cumulative distribution function.

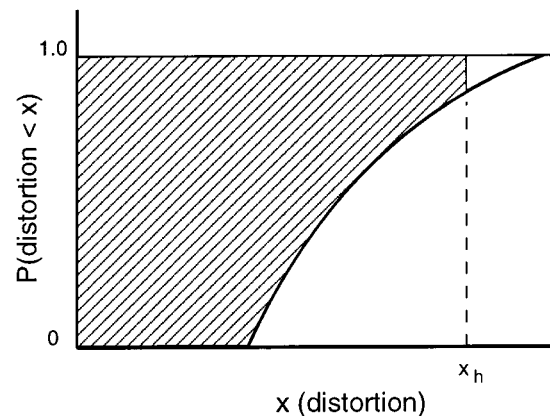


Fig. 5. Shaded area equals the fidelity criterion J_{Φ} with $W(x) = u(-x + x_h)$.

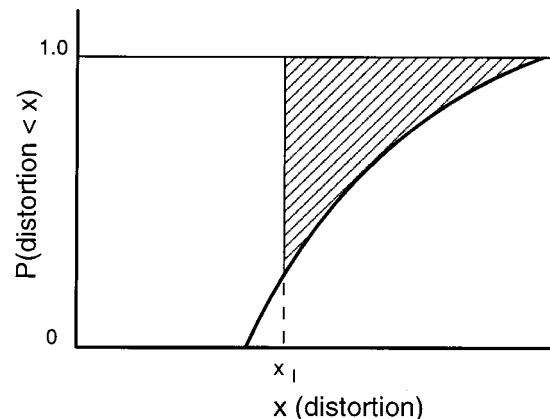


Fig. 6. Shaded area equals the fidelity criterion J_{Φ} with $W(x) = u(x - x_l)$.

In addition, performance requirements can be incorporated into the optimization by restricting the set of system parameters. An *admissibility function* $A(x)$ is a nonnegative real function which provides a pointwise lower bound constraint for acceptable MSE CDF's. Specifically, we require $F_{\Phi}(x) \geq A(x)$ for all x . Figs. 7 and 8 show two examples of admissibility functions in relation to distribution functions. An admissibility function can be used to capture typical constraints such as upper bounds on

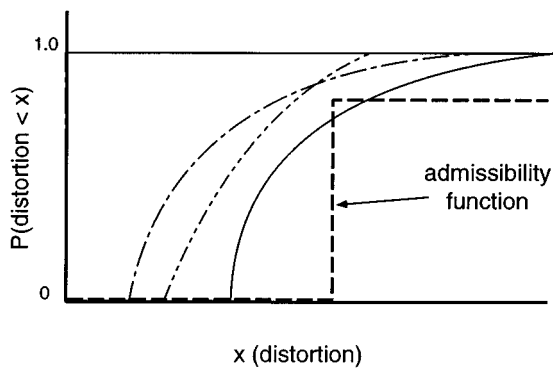


Fig. 7. Admissibility function which places an upper bound on the acceptable probability of exceeding a certain MSE. Two of the three cumulative distributions satisfy the constraint.

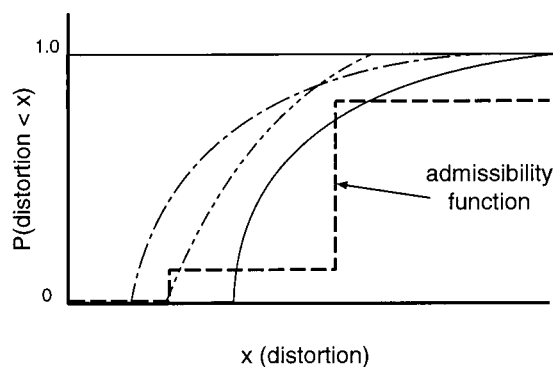


Fig. 8. Admissibility function requiring minimum performance levels for two MSE ranges. Only one of the three distributions satisfies the constraints.

the number of images with high distortion or lower bounds on the number of images with low distortion. For example, Fig. 7 shows an admissibility function where at least 90% of the images can be decoded with MSE less than a given value or equivalently where no more than 10% of the images can be decoded with MSE greater than the same value. Fig. 8 shows an admissibility function which requires a minimum probability for low MSE values in addition to the upper bound on the probability of large MSE values.

The discussion above considers the case of optimizing over a single channel. For systems that operate over a variety of channels, the optimization might need to include a combination of the fidelity criteria from the different channels each with possibly unique weighting and admissibility functions.

We introduce this framework in order to point out how channel code rates were chosen for the current work. For both the hybrid and SPIHT+RCPC/CRC coders, only a finite number of channel codes were available. For each channel code, 1000 trials were run over the channel for which the system was being optimized. The mean decoded MSE (or weighted mean MSE) was calculated for each code rate and the minimum was used as the optimal code. For both cases, the admissibility function was $A(x) = 0.95u(x - 818.6)$ which requires that no more than 5% of the images have an MSE greater than 818.6 (818.6 corresponds to PSNR = 19 dB). For the mean decoded MSE, the weighting function was unity for all x . For the weighted mean MSE, we used the same type of weighting

function as above in Example (ii) with $x_h = 818.6$ and $C = 0$. This function was chosen using data collected from perceptual recognition experiments by Serrano *et al.* [18] in which no successful recognition occurred for images with PSNR < 19 dB. For both algorithms, it was found that the optimal available code was the same for both the weighted and standard mean decoded MSE cases. In practice the admissibility and weighting functions can be chosen according to the specific constraints of a particular application.

V. CHANNEL MODEL

The channel model consists of a combination of a packet erasure channel followed by a discrete channel with memory. This model is used to simulate the end-to-end transmission of images or video from a server through a wireline network to a radio base station which broadcasts the data to mobile receivers.

The wireline network (left side of Fig. 9) will suffer packet losses due to a combination of queue overflow, misrouting, and excessive delay (for video). These packet losses may appear to be bursty for a particular source depending on the relative time span of network impairments compared to the source packet transmission rate. Therefore, in addition to the probability of packet erasure, p_{erasure} , the packet erasure model includes a burst length parameter (i.e., the number of consecutive erased packets), N . Within the model, the source output is divided into groups of N packets, and each group is erased with probability p_{erasure} so that the overall erasure rate is p_{erasure} regardless of the burst length. To test the effect of correlated packet erasures, bursts of lengths $N = 1$ (i.e., independent packet erasures) and $N = 10$ were simulated for each erasure rate.

For the wireless portion of the system (right side of Fig. 9), BPSK transmission over a flat-fading Rayleigh channel was simulated using Jakes' [19] channel model. The channel model was selected to accurately simulate fading channels common in mobile wireless environments. With this model, the channel is characterized by two parameters—the average received signal-to-noise ratio $\overline{\text{SNR}}$, which determines the average bit error rate, and the normalized Doppler spread (i.e., the Doppler spread normalized by the data rate), which determines how quickly the channel changes over time.

The flat-fading characteristic of the channel means there is constant gain across the bandwidth of the received signal. Therefore the effect of the channel is a multiplicative gain term on the received signal level. The quantities SNR_r and $\overline{\text{SNR}}$ in (2)–(6) are on a linear scale, whereas in the text they may be reported instead in decibels. From [20], the bit error rate (BER) of the channel can be written in terms of the average SNR as

$$\text{BER} = \frac{1}{2} \left(1 - \sqrt{\frac{\text{SNR}}{1 + \text{SNR}}} \right). \quad (2)$$

As an example, if the average SNR is 10 dB, then the BER is 0.023, while an average SNR of 20 dB results in a BER of 0.0025. Unlike memoryless channels, such as the BSC and AWGN channels, the channel errors in Jakes' model tend to occur in bursts. Therefore another channel statistic of interest is the average burst or fade duration.

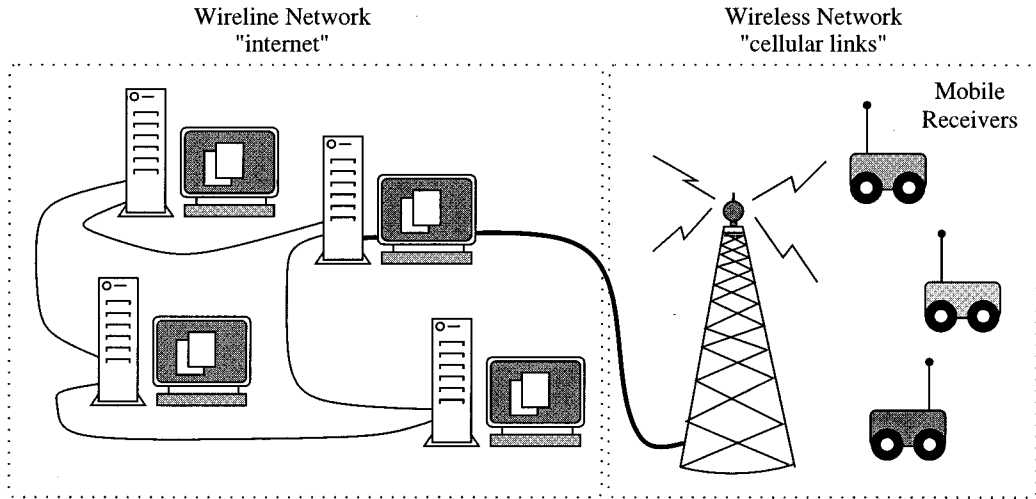


Fig. 9. Schematic of the real world channel. After source coding, packets are sent to the transmitter via the wireline portion of the network (left). Here, packet erasures may occur due to network traffic or queue overflow. Packets which are received by the transmitter are then channel coded and sent to receivers via the wireless network (right). The fading channel here adds bit errors to the transmitted data.

The average burst duration depends on the mobile speed through the normalized Doppler value. Consider an example where the carrier frequency is 900 MHz, the mobile velocity is 4 mi/h, and the data rate is 500 Kbits/s. The maximum Doppler shift is given by

$$f_{\max_Doppler} = f_c \frac{v}{c} \quad (3)$$

where f_c is the carrier frequency, v is the mobile velocity, and c is the speed of light. For this example, we get $f_{\max_Doppler} = 5.38$ Hz. Dividing $f_{\max_Doppler}$ by the data rate gives the normalized Doppler spread, f_D , which is 1.07×10^{-5} for this example. Considering systems where the carrier and data rates are as given in this example and the mobile velocity varies, normalized Doppler values around 10^{-5} represent the low end of the range of interest while values around 2×10^{-4} represent the high end.

The average burst duration also depends on the fade margin (i.e., the necessary received signal level for reliable communication). For BPSK transmission, the probability of error given a received SNR of SNR_r is

$$P_e = Q(\sqrt{2 \cdot \text{SNR}_r}) = \frac{1}{2} \text{erfc}(\sqrt{\text{SNR}_r}). \quad (4)$$

A $P_e = 0.1$ corresponds to $\text{SNR}_r = -0.8556$ dB while $P_e = 0.01$ corresponds to $\text{SNR}_r = 4.3232$ dB. Then, from [21], the average burst duration in bit intervals is given by

$$\bar{\tau} = \frac{e^{\rho^2} - 1}{\rho f_D \sqrt{2\pi}} \quad (5)$$

where ρ is the received amplitude normalized by the RMS amplitude and is computed according to

$$\rho = \sqrt{\frac{\text{SNR}_r}{\text{SNR}}}. \quad (6)$$

With the values $\overline{\text{SNR}} = 10$ dB and $f_D = 10^{-5}$, the average burst duration for $P_e \geq 0.1$ is $\bar{\tau} = 11\,915$ bits and for $P_e \geq 0.01$ is $\bar{\tau} = 23\,832$ bits. As can be seen from these numbers, the low

end of the parameter range of interest corresponds to very long bursts, especially considering the fact that the total transmission rate for the images used in this work at 0.25 b/pixel is 65 536 bits. Channels with longer burst lengths probably require the use of frequency or spatial diversity techniques for reasonable performance.

VI. RESULTS

First, we present results for the special case when there are no packet erasures, in order to demonstrate that the hybrid scheme survives well in this case, and in fact even outperforms its two component encoders under certain channel conditions. Then results for channels with packet erasures and bit errors are presented. In all tests, hard-decision decoding was used in the channel simulation.

To test the robustness of the coders, each was optimized for a fixed channel ($\overline{\text{SNR}} = 13$ dB, $f_D = 10^{-4}$) and then tested over a range of channel conditions. The optimal code rate was chosen using the previously discussed parameter selection method (Section IV). For the hybrid and SPIHT+RCPC/CRC coders, the bits were interleaved using a convolutional interleaver prior to transmission over the fading channel to improve decoding performance. The detailed specifications for the selected codes can be found in the Appendix. All tests were performed using the 512×512 Lena image, and the total transmission rate was fixed at 0.25 b/pixel. Each channel condition was tested with a minimum of 1000 independent trials and as many as 3000 trials on the slowest channels.

In evaluating the performance of the three algorithms, the standard measure of mean decoded MSE may not be sufficient. To improve the analysis, we looked at three characteristics: visual quality, mean decoded MSE, and cumulative distributions of MSE over all trials. The cumulative distributions provide information about the variability of the decoded image quality from one trial to the next. The visual quality, while more difficult to evaluate, is ultimately the performance measure of interest.

TABLE I

DISTORTION VALUES FOR THE THREE ALGORITHMS OVER A RANGE OF CHANNEL CONDITIONS. HYBR AND RCPC WERE OPTIMIZED FOR THE $f_D = 10^{-4}$ AND $\overline{\text{SNR}} = 13$ dB CHANNEL. PZW IS THE PACKETIZED ZEROTREE WAVELET SCHEME [10], RCPC IS THE CODER IN [1], HYBR IS THE COMBINED SCHEME PROPOSED IN THIS PAPER. COLUMN A SHOWS THE MEAN DECODED MSE AND COLUMN B SHOWS A WEIGHTED MEAN USING THE WEIGHTING FUNCTION FROM EXAMPLE (ii): $W(x) = u(-x + 818.6)$

			Average Received SNR(dB)							
			10		12		15		20	
			A	B	A	B	A	B	A	B
Normalized Doppler ($\times 10^{-6}$)	10	HYBR	160.5	158.6	136.6	135.4	115.7	114.5	96.9	96.6
		PZW	458.4	359.4	290.4	251.1	158.7	149.6	92.0	90.0
		RCPC	241.0	165.5	196.0	142.8	146.6	120.8	110.3	100.8
	50	HYBR	181.9	181.7	147.5	147.2	112.7	112.5	91.4	91.1
		PZW	391.6	378.9	303.2	300.4	150.9	150.7	94.6	94.4
		RCPC	244.4	188.1	166.0	139.7	123.1	109.4	93.1	91.2
	100	HYBR	163.3	163.0	122.0	121.7	99.6	99.3	89.0	88.7
		PZW	381.2	379.3	265.1	264.5	158.3	157.5	76.8	76.6
		RCPC	180.1	141.4	126.7	113.3	100.8	97.1	91.4	90.2
	200	HYBR	126.4	126.2	103.8	103.6	92.3	92.0	88.2	87.9
		PZW	784.7	774.6	278.9	278.7	169.6	169.4	87.2	87.0
		RCPC	131.8	119.1	106.6	99.6	96.0	92.7	89.7	89.3



Fig. 10. Images displayed here show the visual effects of loss for the hybrid coder (left) vs. the SPIHT+RCPC/CRC coder (right) at equal MSE (PSNR = 23.5 dB). As shown in [18], the reconstructed images with more localized distortions tend to be preferred over the ones with global blurriness, both in terms of subjective quality and in terms of image content recognizability.

A. Visual Quality

Fig. 10 shows how distortion from a noisy channel is distributed over the image for the hybrid and SPIHT+RCPC/CRC algorithms. Because the hybrid coder groups wavelet coefficient trees together, when losses occur these trees are lost as a single unit. This localizes the region of the image which will be distorted. Trees which are not lost will be decoded with high relative quality (dependent only on the source coding rate). Furthermore, because the regions that were received have good quality, the hybrid algorithm can use the correlation of these neighbors to mitigate the effect of the lost regions. The SPIHT+RCPC/CRC coder distributes the distortion somewhat uniformly over the entire image. The maximum distortion for any pixel may be lower than in the hybrid scheme, but for large total distortion, a spatially distributed error can be perceptually undesirable. This difference in visual quality must be taken into account when interpreting the numerical results. These observations are supported by results of a recent study [18]. In [18], human observers were asked to evaluate a series of images compressed with the SPIHT and PZW algorithms at equal PSNR values. The distortions in the images were similar to those that would result from operation over noisy channels. It was found that the PZW sequences allowed observers to

make responses to objective recognition tasks at lower PSNR than the SPIHT sequences. In addition, subjective rankings on a five-point scale found that PZW images were preferred to SPIHT images at the same PSNR. From these human observer experiments we conclude that the cumulative distribution plots of PSNR are somewhat conservative on the side of underestimating the quality of the hybrid and PZW coders.

B. Mean Decoded MSE

Table I shows the mean decoded MSE values as well as the weighted MSE values for a number of channel conditions which span the range of interest. The largest differences occur for the most severe channels in the upper left section of the table. In these cases, the channels were slow enough that interleaving was not effective and error rates were high due to the low received SNR. The initial packets were often lost for the SPIHT+RCPC/CRC coder which resulted in very high MSE values, and this greatly increased the mean MSE. The hybrid coder has the advantage of being able to continue decoding after error bursts have affected the initial packets (and the initial packets were not more vital to image quality than were later packets).

The hybrid algorithm exhibits superior performance over the SPIHT+RCPC/CRC coder over many channels in terms of mean decoded MSE (Column A in Table I). The weighted distortion (Column B) tends to benefit the SPIHT+RCPC/CRC coder more than the others. In this case, images with very large distortion (above x_h) are not included in the average and therefore do not skew the average. Only the percentage of such images affects this performance measure. Even with this performance measure, the hybrid shows superior performance on the most severe channels (upper left) compared with either of the two other algorithms.

C. Cumulative Distributions of MSE

Further information on performance is provided in Fig. 11(a)–(d). The plots show the cumulative distributions of the decoded MSE for the four channels whose parameter values

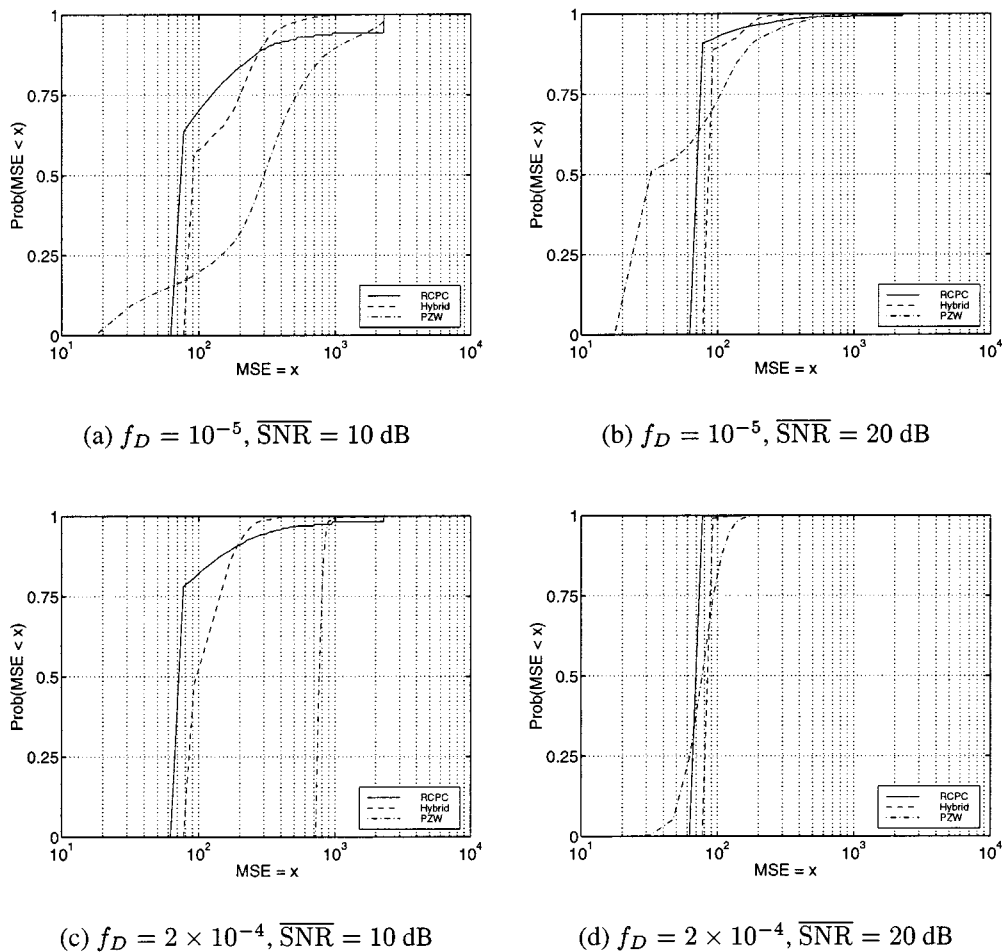


Fig. 11. Cumulative distributions of decoded MSE for the 512×512 Lena image transmitted over Rayleigh fading channels with a transmission rate of 0.25 b/pixel.

are at the corners of Table I. Curves closer to the left and top sides of the plot have better performance in this type of graph.

The initial packet losses for the SPIHT+RCPC/CRC coder previously mentioned are visible as a relatively high tail for the distribution in Fig. 11(a). Even though less than 10% of the decoded images have these large distortions (high MSE), the very large MSE values have a considerable effect on the mean MSE. Notice the performance generally improves for higher received SNR's (due to fewer channel errors) as well as for faster channels because the interleaver is more effective. By contrast, the PZW coder performance degrades for faster channels because the errors are less bursty, so more packets are lost for a given average error rate. In addition, because the effective visual performance in the high distortion regions is worse for the SPIHT+RCPC/CRC algorithm (Figs. 10 and 12), we see that the overall performance of the hybrid is superior over the channels of interest.

D. With Packet Erasure

The next set of results includes the effects of the packet erasure channel. Table II compares the algorithms' performances over varying packet erasure channels (the wireless channel

parameters are fixed). Notice that the SPIHT+RCPC/CRC performs well for a packet burst length of 10 (using either mean MSE or weighted mean MSE). The reason for the improved performance on bursty erasure channels is due to the higher likelihood of the first packet erasure occurring late in the transmission for a fixed erasure rate. Both PZW and the hybrid coders show robust performance over the varying channels, but the hybrid is able to produce higher quality images on average. Fig. 13(a)–(d) shows the cumulative distributions of performance for these different packet erasure conditions. Numerically the hybrid performs competitively with the SPIHT+RCPC/CRC. After considering the fact that in the high distortion regime the hybrid visual performance is better than its numerical results might suggest (based on results in [18]), we see that the hybrid coder is superior over this range of channels. Visual results over two particular channels are shown in Fig. 12. We see that the SPIHT+RCPC/CRC coder has strong performance over the channel dominated by bit errors, but performance is significantly degraded on the channel dominated by packet erasures. The PZW and hybrid scheme show consistent performance over both channels. But because of the additional error-correction capability, the hybrid decodes images at a higher quality on average.

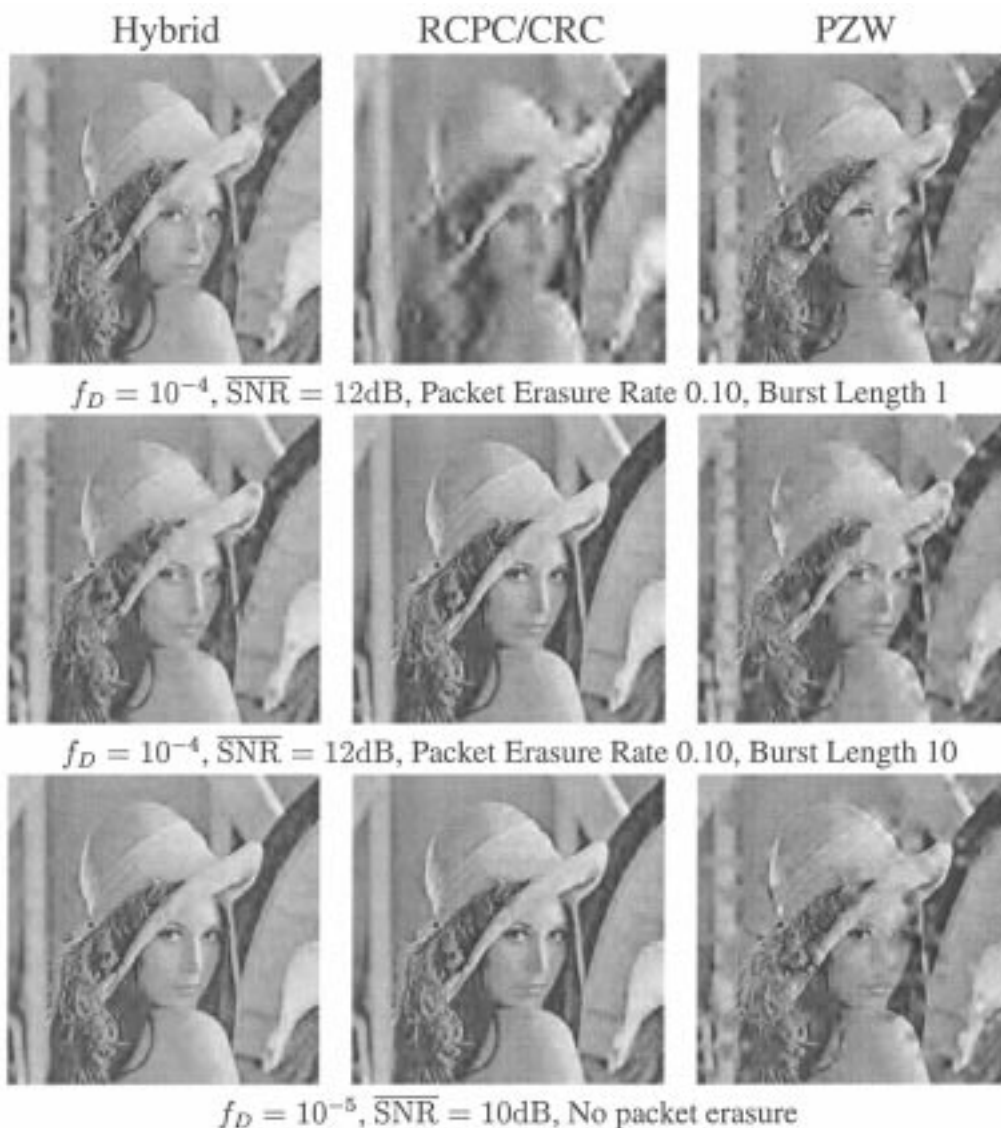


Fig. 12. Images displayed here show the median quality for the three algorithms under different channel conditions. The channel used for the top row of images was dominated by uniform packet erasures. The second row includes representative images over a channel with packet erasures occurring in bursts. The channel for the bottom row of images has no packet erasures but has a higher probability of bit errors, and in long bursts. Overall transmission rate was 0.25 b/pixel.

TABLE II

DISTORTION VALUES FOR THE THREE ALGORITHMS OVER A RANGE OF CHANNEL CONDITIONS. FOR THESE CHANNELS, THE FADING PARAMETERS WERE HELD CONSTANT WHILE THE PACKET ERASURE CHANNEL PARAMETERS WERE VARIED. HYBR AND RCPC WERE OPTIMIZED FOR THE $f_D = 10^{-4}$ AND $\overline{\text{SNR}} = 13$ dB CHANNEL. PZW IS THE PACKETIZED ZEROTREE WAVELET SCHEME [10]. RCPC IS THE CODER IN [1]. HYBR IS THE COMBINED SCHEME PROPOSED IN THIS PAPER. COLUMN A SHOWS THE MEAN DECODED MSE AND COLUMN B SHOWS A WEIGHTED MEAN USING THE WEIGHTING FUNCTION FROM EXAMPLE (ii): $W(x) = u(-x + 818.6)$

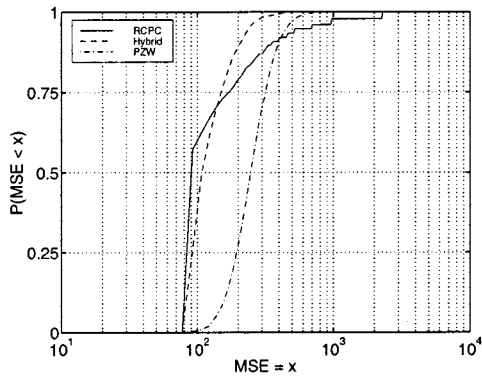
		Burst Length				
		1		10		
		A	B	A	B	
Packet Erasure Rate	0.01	HYBR	137.6	137.3	136.6	136.4
		PZW	273.0	272.6	274.0	273.6
		RCPC	195.2	167.9	137.5	121.5
	0.10	HYBR	206.9	206.7	200.5	200.1
		PZW	345.7	345.3	340.8	340.2
		RCPC	614.5	439.1	208.0	174.2

VII. CONCLUSION

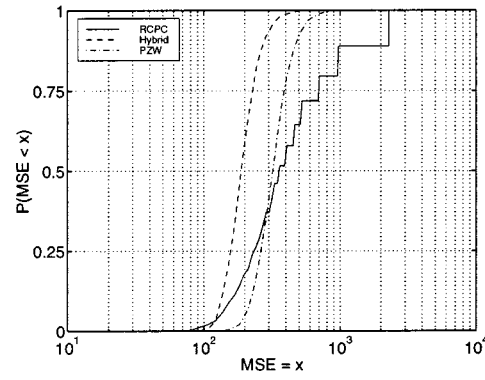
In many applications, the system must operate in a highly variable environment. In these cases, the source and channel coders must be able to handle a large range of potential conditions. In addition, severe channels can lead to large variations in decoded image quality over different trials, making it difficult to decisively conclude which coding method is superior. Using the mean decoded MSE as well as cumulative distribution plots and the visual results obtained by this research, we conclude that the hybrid coder performs competitively across all channel conditions and degrades more gracefully under the most severe conditions.

APPENDIX
CODE PARAMETERS

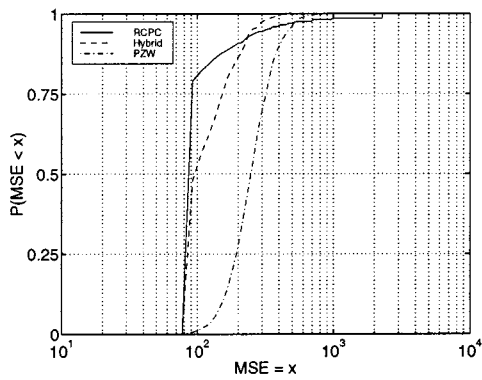
This appendix lists the channel code parameters used in this paper. All packets consisted of 384 source bits. The polyno-



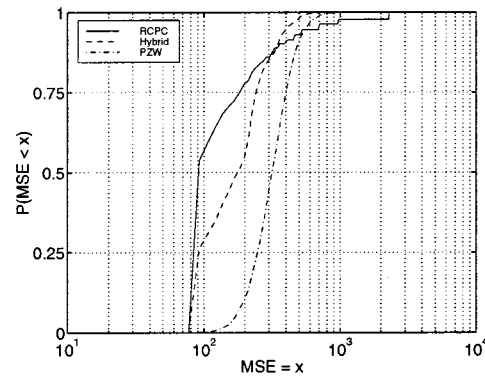
(a) Packet Erasure Rate 0.01, Burst Length 1



(b) Packet Erasure Rate 0.10, Burst Length 1



(c) Packet Erasure Rate 0.01, Burst Length 10



(d) Packet Erasure Rate 0.10, Burst Length 10

Fig. 13. Cumulative distributions of decoded MSE for the 512×512 Lena image over channels with varying packet erasure parameters and fading parameters $f_D = 10^{-4}$, SNR = 12 dB. (Total transmission rate 0.25 b/pixel.)

TABLE III
RCPC CODES

Rate	Mother Code	Puncturing Matrix	Rate	Mother Code	Puncturing Matrix
4/9	155	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$	4/13	155	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$
	123	$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$		123	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$
	137	$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$		137	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$
	147	$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$		147	$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$

mials below will be expressed in octal notation (e.g., octal 13 is 001 011 in binary which translates to the polynomial $X^3 + X + 1$).

All codes used a 16 bit CRC defined by the polynomial 254465. All RCPC codes in this paper were constructed from memory 6 mother codes, and each packet was terminated with enough zero bits to flush the state of the convolutional coder (i.e., 6 bits in this case). The search for the correct path in the list-Viterbi algorithm was terminated after 100 candidates as in [2]. However, note that limiting the search depth to 10 candidates gives almost the same performance. The hybrid coder used the rate 4/9 RCPC code and the SPIHT+RCPC/CRC used the rate 4/13 RCPC code. The puncturing matrices and mother

codes are listed in Table III. A convolutional bit interleaver was used for the hybrid and SPIHT+RCPC/CRC coders for transmission over the fading portion of the channel. This type of interleaver operates in a more continuous fashion compared to a block interleaver, so it can be more easily matched to any total transmission rate. There is a penalty on the order of n^2 bits for an interleaver of depth n as initial zero bits in the memory are flushed. Increasing the interleaver depth tends to make the de-interleaved errors more uniform (helping the decoding performance of the RCPC codes), but there is a penalty in the total number of source bits received for any total transmission rate. The interleaving depth for the hybrid coder was 70 and for the SPIHT+RCPC/CRC code was 70.

ACKNOWLEDGMENT

The authors would like to thank N. Serrano and D. Schilling for their study of the visual performance of the algorithms, and E. Riskin and the reviewers for their helpful remarks.

REFERENCES

- [1] P. G. Sherwood and K. Zeger, "Progressive image coding on noisy channels," in *Proc. DCC '97*, J. A. Storer and M. Cohn, Eds., 1997, pp. 72–81.
- [2] ———, "Progressive image coding for noisy channels," *IEEE Signal Processing Lett.*, vol. 4, pp. 189–191, July 1997.
- [3] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits, Syst. Video Technol.*, vol. 6, pp. 243–250, June 1996.
- [4] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, pp. 3445–3462, Dec. 1993.
- [5] Z. Xiong, B.-J. Kim, and W. A. Pearlman, "Progressive video coding for noisy channels," in *Proc. ICIP 98*, vol. 1, 1998, pp. 334–337.
- [6] V. Chande, H. Jafarkhani, and N. Farvardin, "Joint source-channel coding of images for channels with feedback," in *Proc. Information Theory Workshop*, Feb. 1998, pp. 50–51.
- [7] I. Kozintsev, J. Chou, and K. Ramchandran, "Image transmission using arithmetic coding based continuous error detection," in *Proc. DCC '98*, J. Storer and M. Cohn, Eds., 1998, pp. 339–348.
- [8] H. Man, F. Kossentini, and M. J. Smith, "A class of EZW image coders for noisy channels," in *Proc. ICIP 97*, vol. 3, 1997, pp. 90–93.
- [9] V. Crump and T. Fischer, "Intraframe low bitrate video coding robust to packet erasure," in *Proc. DCC '97*, J. Storer and M. Cohn, Eds., 1997, p. 432.
- [10] J. Rogers and P. Cosman, "Robust wavelet zerotree image compression with fixed-length packetization," in *Proc. DCC '98*, J. Storer and M. Cohn, Eds., 1998, pp. 418–427.
- [11] ———, "Wavelet zerotree image compression with packetization," *IEEE Signal Processing Lett.*, vol. 5, pp. 105–107, May 1998.
- [12] C. Creusere, "A new method of robust image compression based on the embedded zerotree wavelet algorithm," *IEEE Trans. Image Processing*, vol. 6, pp. 1436–1442, Oct. 1997.
- [13] P. G. Sherwood and K. Zeger, "Error protection for progressive image transmission over memoryless and fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 1555–1559, Dec. 1998.
- [14] B. S. Srinivas, E. A. Riskin, R. Ladner, and M. Azizoglu, "Progressive image transmission on a channel with memory," in *Proc. 33rd Annu. Allerton Conf.*, 1995, pp. 265–274.
- [15] P. G. Sherwood and K. Zeger, "Macroscopic multistage image compression for robust transmission over noisy channels," *Proc. SPIE*, vol. 3653, pp. 73–83, Jan. 1999.
- [16] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. Commun.*, vol. 36, pp. 389–400, Apr. 1988.
- [17] G. Zemor and G. D. Cohen, "The threshold probability of a code," *IEEE Trans. Inform. Theory*, vol. 41, pp. 469–477, Mar. 1995.
- [18] N. Serrano, D. Schilling, and P. Cosman, "Quality evaluation for robust wavelet zerotree image coders," in *Proc. 2nd Annu. UCSD Conf. Wireless Communications*, Mar. 1999, pp. 128–134.
- [19] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.
- [20] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.
- [21] T. S. Rappaport, *Wireless Communications: Principles & Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1996.



Pamela C. Cosman (S'90–M'93) received the B.S. degree with honors in electrical engineering from the California Institute of Technology, Pasadena, in 1987, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1989 and 1993, respectively.

She was an NSF Postdoctoral Fellow at Stanford University and a Visiting Professor at the University of Minnesota, Minneapolis, from 1993 to 1995. Since July 1995, she has been on the faculty of the Department of Electrical and Computer Engineering, University of California, San Diego, where she is currently an Associate Professor. Her research interests are in the areas of data compression and image processing.

Dr. Cosman is the recipient of the ECE Departmental Graduate Teaching Award (1996), a CAREER Award from the National Science Foundation (1996–1999), and a Powell Faculty Fellowship (1997–1998). She is an Associate Editor of the *IEEE COMMUNICATIONS LETTERS* and a member of Tau Beta Pi and Sigma Xi.



Jon K. Rogers received the B.S. degree in physics from California Polytechnic State University, San Luis Obispo, in 1992, and the M.S. degree in imaging science from the Rochester Institute of Technology, Rochester, NY, in 1994. He did graduate work in the Department of Electrical and Computer Engineering, University of California, San Diego, where his research focused on image compression and noisy channel transmission.

He is currently a Software Engineer with Zing Networks, San Francisco, CA.



P. Greg Sherwood (S'88–M'90) received the S.B. degree in physics and the S.B. and S.M. degrees in electrical engineering in 1990, all from the Massachusetts Institute of Technology, Cambridge, and the Ph.D. degree in electrical engineering from the University of Illinois, Urbana-Champaign, in 2000.

From 1990 to 1993, he was a Software/Systems Engineer with Engineering Research Associates, Vienna, VA. From 1997 to 1999, he was with TRW Avionics Systems Division working on voice and data communications systems. In 1999, he joined

Packet Video, San Diego, CA, where he works on multimedia applications for embedded and wireless systems.



Kenneth Zeger (S'85–M'90–SM'95–F'00) was born in Boston, MA, in 1963. He received the S.B. and S.M. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1984, and the M.A. degree in mathematics and the Ph.D. degree in electrical engineering from the University of California, Santa Barbara, in 1989 and 1990, respectively.

He was an Assistant Professor of electrical engineering at the University of Hawaii, Honolulu, from 1990 to 1992. He was an Assistant Professor with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois, Urbana-Champaign, from 1992 to 1995, and an Associate Professor from 1995 to 1996. He has been with the Department of Electrical and Computer Engineering, University of California, San Diego, as an Associate Professor from 1996 to 1998, and as a Professor from 1998 to present.

Dr. Zeger received the NSF Presidential Young Investigator Award in 1991. He served as Associate Editor At-Large for the *IEEE TRANSACTIONS ON INFORMATION THEORY* from 1995 to 1998, and is serving as a Member of the Board of Governors of the IEEE Information Theory Society during 1998–2000.