Real-time Unequal Error Protection for Distortion-Optimal Progressive Image Transmission

Vladimir Stanković, Youssef Charfi, Raouf Hamzaoui
University of Konstanz
Department of Computer and Information Science
D-78457 Konstanz
stankovi,charfi,hamzaoui@fmi.uni-konstanz.de

Zixiang Xiong
Texas A &M University
College Station, TX 77843
zx@lena.tamu.edu

***Abstract*— For optimal progressive transmission of an embedded image code over a noisy channel, we consider an unequal error protection strategy that minimizes the average of the expected distortion over a set of intermediate rates. In contrast to previous work, we find a near-optimal solution in real-time. For a binary symmetric channel, two state-of-the-art source coders (SPIHT and JPEG2000), and a rate-compatible punctured turbo coder as a channel coder, we compare our solution to the strategy that optimizes the end-to-end performance.**

I. Introduction

A progressive image coder generates a bitstream such that the reconstruction quality at the receiver improves as more and more bits are decoded. This is an important property in many applications. In Internet browsing, for example, it allows the receiver to stop downloading an image as soon as an acceptable fidelity is reached. The new generation wireless multimedia communication systems, which provide access to the Internet, should enable real-time progressive decoding.

Embedded coders form an important class of progressive coders. An embedded coder is such that for any two bitstreams generated by the encoder, the shorter one is a prefix of the longer one. In addition to being progressive, an embedded coder can compress the image once at a high transmission rate and truncate the compressed file to obtain the image code at any lower transmission rate. Two of the most popular embedded coders are the Set Partitioning in Hierarchical Trees (SPIHT) [1] coder and JPEG2000 [2].

When the transmission channel is prone to errors, the communication system must include a protection strategy, which should be adapted to the nature of the channel errors. One powerful system for the protection of embedded codes against errors in memoryless noisy channels was proposed by Sherwood and Zeger [3]. The system uses a channel code consisting of a concatenation of an error detection code and an error correction code taken from a set of rate-compatible convolutional codes. Whenever an error is detected, the decoding is stopped, and the image is reconstructed from the correctly received packets. To optimize the progressive performance of this system, Sherwood, Tian, and Zeger [4] proposed to find an unequal error protection solution that minimizes the average of the expected distortion over the set of appropriate intermediate transmission rates. In this way, the performance of the system is considered not only at the target transmission rate, but also at many intermediate rates. They also devised a dynamic programming algorithm for computing an optimal solution. The time complexity of the algorithm is quadratic in the target transmission rate when the number of source bits per packet is fixed. Thus, the computation of a solution in not possible in real-time. Moreover, for the more convenient system where the number of source bits in a packet is variable, while the length of the channel codewords is fixed, the time complexity is exponential in the number of packets. An alternative is to maximize the average of the expected number of correctly received source bits [4], [5], which is a reasonable approach for an embedded source code. Now, an optimal protection can be found with a linear-time algorithm [6]. Moreover, the solution is independent of the image and the source coder. Thus, it need not be specified to the decoder.

In this paper, we address the problem of efficiently minimizing the distortion-based progressive performance of [4] in the context of fixed-length channel codewords. We consider a binary symmetric channel, which is often used to model wireless links. We show that the fast local search algorithm introduced in [7] for minimizing the end-to-end performance of the system can be successfully adapted to the more general progressive measure of [4]. In contrast to [4], our solution has a linear-time worst-case complexity and can be computed in real-time. Experimental results for a binary symmetric channel and two embedded source coders (SPIHT and JPEG2000) show that at most intermediate transmission rates, especially at the lowest ones, the performance of the local search solution that minimizes the average of the expected distortion is significantly better than that of the solution that minimizes the expected distortion at the target rate, whereas the performance of both solutions is similar at the target rate.

II. TERMINOLOGY

We consider a joint source-channel system where the source coder is a progressive coder and where the channel coder uses a finite number of channel codes with error detection and error correction capability. The channel encoder transforms the information bitstream into a sequence of channel codewords of fixed length L. These codewords (packets) are sent over a binary symmetric channel. If a packet is correctly decoded, then the next packet is considered. Otherwise, the decoding is stopped, and the image is reconstructed from the correctly decoded packets. We assume that all errors can be detected. Given m channel codes, let $\mathcal{R}$ be the set of corresponding code rates $r_1 < \cdots < r_m$ with $p(r_1) < \cdots < p(r_m)$, where $p(r_i)$ denotes the probability of a decoding error in a packet protected with r_i. An N*-packet error protection scheme* (EPS) $R = (r_{k_1}, \ldots, r_{k_N})$ assigns to each packet i, $i = 1, \ldots, N$, a channel code rate $r_{k_i} \in \mathcal{R}$. Then for $i = 1, \ldots, N-1$, $P_i(R) = \prod_{j=1}^{i}(1 - p(r_{k_j}))p(r_{k_{i+1}})$ is the probability that no errors occur in the first i packets with an error in the next one, $P_0(R) = p(r_{k_1})$ is the probability of an error in the first packet, and $P_N(R) = \prod_{j=1}^{N}(1 - p(r_{k_j}))$ is the probability that all N packets are correctly decoded. The expected distortion associated to an N-packet EPS R is

$$E_N[d](R) = \sum_{i=0}^{N} P_i(R) d_i(R), \tag{1}$$

where $d_0(R) = d_0$ is a constant, and for $i \geq 1$, $d_i(R)$ is the reconstruction error using the first i packets. An N-packet EPS $R \in \mathcal{R}^N$ that minimizes (1) is called *distortion optimal*.

The expected number of correctly decoded source bits for an N-packet EPS R is

$$E_N[r](R) = \sum_{i=0}^{N} P_i(R) V_i(R), \tag{2}$$

where $V_0(R) = 0$ and for $i \geq 1$, $V_i(R)$ is the number of source bits in the first i packets. An N-packet EPS that maximizes (2) is called *rate optimal*. A rate-optimal solution can be computed in linear time [8]. Moreover, a rate-optimal solution $(r_{k_1}, \ldots, r_{k_N})$ satisfies $r_{k_1} \leq \cdots \leq r_{k_N}$ [8].

To measure the progressive performance of an N-packet EPS $R = (r_{k_1}, \ldots, r_{k_N})$, one may compute the average of the expected performance over all intermediate rates [4]. This is

$$L_N[\phi](r_{k_1}, \ldots, r_{k_N}) = \frac{1}{N} \sum_{n=1}^{N} E_n[\phi](r_{k_1}, \ldots, r_{k_n}), \tag{3}$$

where ϕ may be the distortion or the source rate. An EPS that minimizes the average expected distortion is called *progressive distortion optimal (PDO)*, whereas an EPS that maximizes the average expected source rate is called *progressive rate optimal (PRO)*. In [6], Stanković and Hamzaoui proposed an algorithm that finds a progressive rate-optimal solution in $O(N)$ time.

A distortion-optimal (respectively progressive distortion-optimal) solution such that $r_{k_1} \leq \cdots \leq r_{k_N}$ is called *constrained distortion optimal (CDO)* (respectively *progressive constrained distortion optimal (PCDO)*). This constraint reduces the number of candidates from m^N to $\binom{m+N-1}{N}$.

Finally, in situations where one wants to emphasize the progressive performance at some particular intermediate rates, one can use the weighted cost function [4]

$$L_{(N;\omega_1,\ldots,\omega_N)}[\phi](R) = \frac{1}{N} \sum_{n=1}^{N} \omega_n E_n[\phi](r_{k_1}, \ldots, r_{k_n}), \tag{4}$$

where the weights ω_n, $n = 1, \ldots, N$, are chosen in $[0, 1]$. Note that the cost function (3) is simply $L_{(N;1,\ldots,1)}[\phi]$, while the cost functions (1) and (2) are obtained from $L_{(N;0,\ldots,0,1)}[\phi]$.

III. PROGRESSIVE DISTORTION-OPTIMAL SOLUTION

In this section, we show that the main results of [7], obtained for the cost functions (1) and (2), can be extended to the more general cost function (3) (and also to (4)). The basic result is that under reasonable conditions a progressive distortion-optimal solution must use more protection bits than a progressive rate-optimal solution.

Proposition 1: Let f be the operational distortion-rate function of the source coder. Suppose that f is nonincreasing and convex. Let T^* be an N-packet progressive distortion-optimal EPS and let R^* be an N-packet progressive rate-optimal EPS. Let $V_N(R)$ denote the number of source bits protected with R. Then $V_N(T^*) \leq V_N(R^*)$, and the inequality is strict if T^* is not progressive rate optimal.

Proof: Let $R = (r_{k_1}, \ldots, r_{k_N})$ be an N-packet EPS. For $n \leq N$, let $R_n = (r_{k_1}, \ldots, r_{k_n})$. Then we have

$$\begin{aligned} L_N[\phi](R) &= \frac{1}{N} \sum_{n=1}^{N} \sum_{i=0}^{n} P_i(R_n) F_i(R_n) \\ &= \sum_{i=0}^{N} \Omega_i(R) F_i(R), \end{aligned}$$

where ϕ denotes the distortion (respectively the source rate), $F_i(R)$ is $d_i(R)$ (respectively $V_i(R)$), and

$$\Omega_i(R) = \begin{cases} p(r_{k_1}), & \text{if } i = 0; \\ \sum_{i \leq n \leq N} \frac{1}{N} P_i(R_n), & \text{otherwise.} \end{cases}$$

The proof follows then from Proposition 1 (ii) of [7] with the observation that $\sum_{i=0}^{N} \Omega_i(R) = 1$.

■

Using Proposition 1, the following algorithm efficiently computes a progressive distortion-optimal EPS T^* by discarding all candidates that use fewer protection bits than a progressive rate-optimal solution.

Algorithm PDO

Suppose that the operational distortion-rate function is non-increasing and convex. Let R^* be a progressive rate-optimal N-packet EPS. Let $A_{R^*} = \{R_1, \ldots, R_i, \ldots, R_{n_R}\} \subset \mathcal{R}^N$ be the finite set of all N-packet EPS's R that use more protection bits than R^*, and such that $V_N(R_{i+1}) \leq V_N(R_i)$.

1. Set $i = 1$. Use the algorithm of [6] to compute R^*.

2. Let R_i be the ith EPS in A_{R^*}. If $L_N[d](R_i) < L_N[d](T^*)$, set $T^* = R_i$.

3. If $d_N(R_i) > L_N[d](T^*)$, stop.

4. Set $i = i + 1$. If $i > n_R$, stop. Otherwise, go to Step 2.

Although Algorithm PDO significantly reduces the complexity of the optimization, the computation time may be unacceptable for real-time applications. However, one can quickly find a near-optimal progressive distortion-optimal solution by adapting the local search algorithm of [7] to the cost function (3). This gives the following local search algorithm, which starts from a progressive rate-optimal solution and converges to a local minimum of (3).

Algorithm PLS

1. Set $k = 1, l = 1$, and $n = 0$. Use the algorithm of [6] to compute a progressive rate-optimal N-packet EPS R_n.

2. Let r be the kth highest rate used by R_n. Let j be the index of the first packet that R_n protects with r. If $r = r_1$, stop. Otherwise, let $r_c \in \mathcal{R}$ be the lth highest rate smaller than r and define R_c to be the EPS obtained from R_n by protecting packet j with r_c.

3. If $L_N[d](R_c) < L_N[d](R_n)$, set $R_{n+1} = R_c, n = n + 1$, and go to Step 2.

4. If $j \neq 1$ and r_c is greater than the rate of packet $j - 1$, set $l = l + 1$. If $j \neq 1$ and r_c is equal to the rate of packet $j - 1$, set $l = 1$ and $k = k + 1$. If $j = 1$ and $r_c \neq r_1$, set $l = l + 1$. If $j = 1$ and $r_c = r_1$, stop.

5. Go to Step 2.

It is easy to check that the worst-case time complexity of the above algorithm is $O(Nm)$.

IV. Results

We present results for the transmission of the standard 8 bits per pixel (bpp) 512 x 512 Lenna image over a binary symmetric channel. We obtained similar results with the 512 x 512 Goldhill image. We used two state-of-the-art embedded source coders: the SPIHT coder [1] and the Kakadu implementation of JPEG2000 in the distortion scalable mode [2]. The channel coder was a concatenation of a 32-bit CRC coder and a rate-compatible punctured turbo (RCPT) coder [9]. The generator polynomial of the CRC code was $(32, 26, 23, 22, 16, 12, 11, 10, 8, 7, 5, 4, 2, 1, 0)$. The turbo coder consisted of two identical recursive systematic convolutional encoders with memory length 4 and generators $(31, 27)$ (octal). The mother code was $20/60 = 1/3$, and the puncturing rate was 20, yielding various channel code rates. The length of a packet was equal to $L = 2048$ bits, consisting of a variable number of source bits, 32 CRC bits, 4 bits to set the turbo encoder into a state of all zeroes, and protection bits. We used iterative maximum a posteriori decoding, which was stopped if no correct sequence was found after 20 iterations. The bit error rate (BER) of the binary symmetric channel was 0.1. The probability of a packet decoding error for each code rate was computed with 50000 Monte Carlo simulations. The set $\mathcal{R}$ consisted of the four channel code rates $r_1 = \frac{20}{56}, r_2 = \frac{20}{52}, r_3 = \frac{20}{50}$, and $r_4 = \frac{20}{48}$ [7].

Fig. 1 shows the peak-signal-to-noise ratio (PSNR) of the expected mean squared error (MSE) of a progressive rate-optimal (PRO) solution [6], a progressive constrained distortion-optimal solution (PCDO) computed with Algorithm PDO, and a solution computed with Algorithm PLS (we call this solution progressive local search (PLS) solution) as a function of the intermediate transmission rate. We used the SPIHT coder. The target transmission rate was 0.5 bpp. Thus, the cost function (3) was minimized for $N = 512 \times 512 \times \frac{0.5}{2048} = 64$ packets. Fig. 2 shows the results of the same experiments for JPEG2000. Both the PCDO and the PLS solutions outperformed the PRO solution at many intermediate rates, including the target rate. For example, for JPEG2000 the PSNR gain of PCDO and PLS over PRO at 0.5 bpp was 0.75 dB and 0.64 dB, respectively. The difference between the PCDO and the PLS solutions was always negligible, showing that our local search strategy was almost optimal. This result is stressed in Fig. 3, which shows the difference in L_N[MSE], the average expected MSE, between PLS and PCDO for several target transmission rates. The singularity for JPEG2000 at target transmission rate 0.15625 bpp ($N = 20$) was due to the fact that for this rate the PLS algorithm did not improve the PRO solution.

Table I compares the time complexity of the different algorithms for the SPIHT coder. For each algorithm, the optimization was done for four target transmission rates. The CPU time was measured on a 270 MHz MIPS R12000 processor of an SGI Origin200 server with a main memory size of 1.5 Gbytes. The programs were written in C and compiled with the -O3 optimization option. For PCDO and PLS, the reported time includes the time spent to read the distortion-rate points from a file. This step can be avoided if a parametric model is used for the operational distortion-rate function of the source coder [10]. Note that PRO does not require the distortion-rate function because the solution is independent of both the image and the source coder [6]. The results show that a PLS solution can be computed in real-time. Table II shows the corresponding allocation of the channel code rates. Note that code rate 20/56 was never selected.

Fig. 4 shows the difference in the expected MSE between the constrained distortion-optimal (CDO) and the PLS solutions at the possible intermediate rates. The source coder was SPIHT and the target transmission rate was 1 bpp. Fig. 5 presents results of the same simulations for JPEG2000. The results show that the PLS solutions had an equal or a better performance

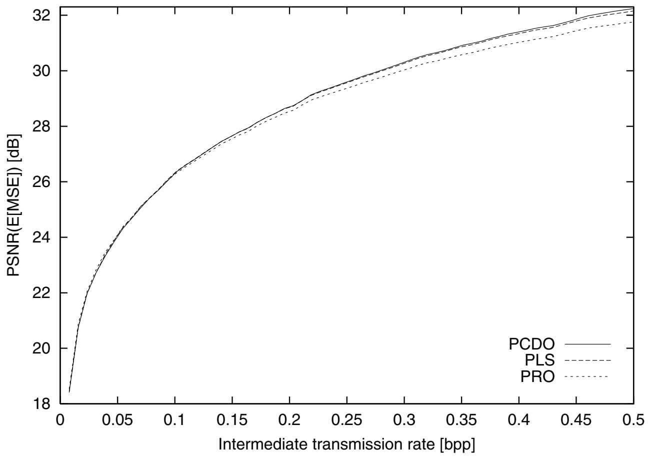

Fig. 1. PSNR of the expected MSE at all possible intermediate rates for a progressive constrained distortion-optimal (PCDO) solution, a progressive local search (PLS) solution, and a progressive rate-optimal (PRO) solution. The source coder is SPIHT, and the target transmission rate is 0.5 bpp.

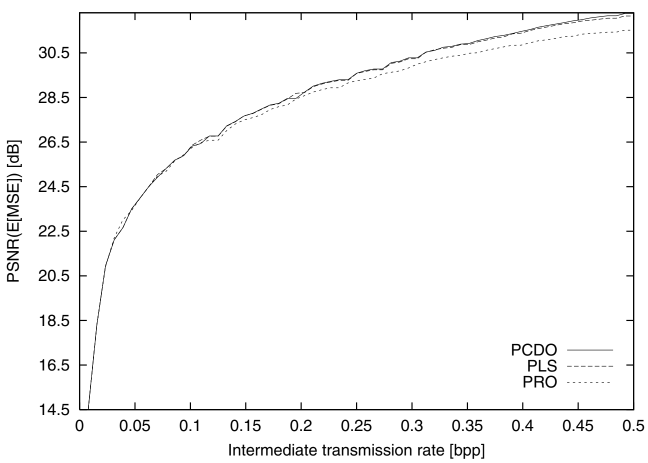

Fig. 2. PSNR of the expected MSE for a progressive constrained distortion-optimal (PCDO) solution, a progressive local search (PLS) solution, and a progressive rate-optimal (PRO) solution. The source coder is JPEG 2000, and the target transmission rate is 0.5 bpp.

at most of the intermediate rates and a slightly worse performance at rates close to the target rate. The superiority of PLS was greatest at the low intermediate rates. Note that the vertical axis was truncated at 50 for clarity of display.

Suppose now that the target transmission rate is 1 bpp and that the receiver wants to reconstruct the image at the intermediate rates 0.25, 0.5, 0.75, and 1.0 bpp only. Then, the local search algorithm should be applied to the weighted cost function (4) with $\omega_i = 1$ if $i \in \{32, 64, 96, 128\}$ and $\omega_i = 0$, otherwise. Table III gives the expected MSE of this solution and compares it to that of a constrained distortion-optimal solution and a progressive local search solution.

We obtained similar results for other channel bit error rates and when a rate-compatible punctured convolutional coder was used instead of the RCPT coder.

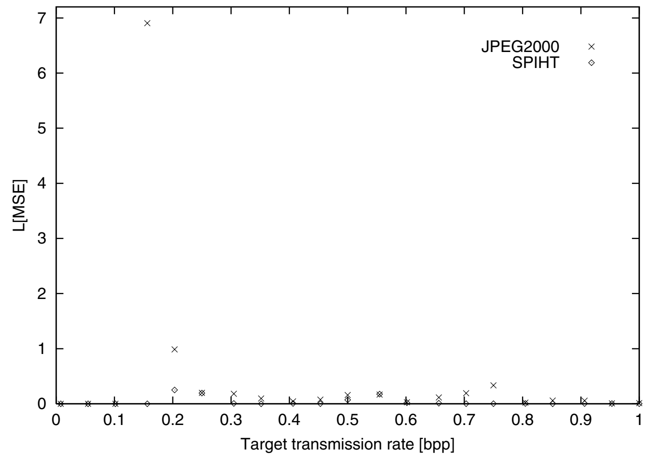

Fig. 3. Difference in the average expected MSE between a progressive local search solution and a progressive constrained distortion-optimal solution for various target transmission rates.

TABLE I

CPU TIME IN SECONDS AT VARIOUS TARGET TRANSMISSION RATES FOR A PROGRESSIVE RATE-OPTIMAL SOLUTION (PRO), A PROGRESSIVE CONSTRAINED DISTORTION-OPTIMAL SOLUTION (PCDO), AND A PROGRESSIVE LOCAL SEARCH SOLUTION (PLS). THE SOURCE CODER IS SPIHT.

Trans. rate (bpp)	PRO Time (s)	PCDO Time (s)	PLS Time (s)
0.25	$<$ 0.01	6.47	0.04
0.5	$<$ 0.01	225.63	0.08
0.75	$<$ 0.01	1774.61	0.3
1.0	$<$ 0.01	6915.3	0.36

V. CONCLUSIONS

We showed that determining an unequal error protection that minimizes the average expected distortion of a source-channel coding system over a set of intermediate rates is feasible in real-time with a local search technique.

At all intermediate rates, the solution has the same or a bet-

TABLE II

CODE RATE ALLOCATION OF TABLE I. THE NOTATION (a, b, c) MEANS THAT CODE RATES 20/52, 20/50, 20/48 WERE SELECTED a, b, c TIMES, RESPECTIVELY.

Rate (bpp)	PRO	PCDO	PLS
0.25	(0,7,25)	(0,14,18)	(2,10,20)
0.5	(0,39,25)	(4,36,24)	(7,37,20)
0.75	(10,61,25)	(30,44,22)	(33,42,21)
1.0	(42,61,25)	(56,47,25)	(61,45,22)

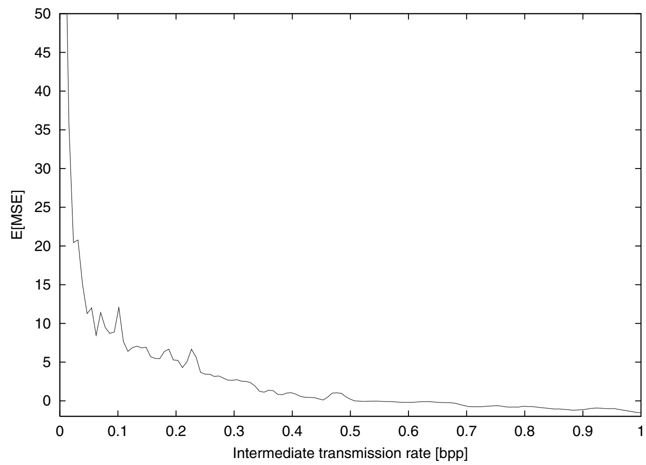

Fig. 4. Difference in expected MSE between CDO and PLS for SPIHT. The target transmission rate is 1 bpp.

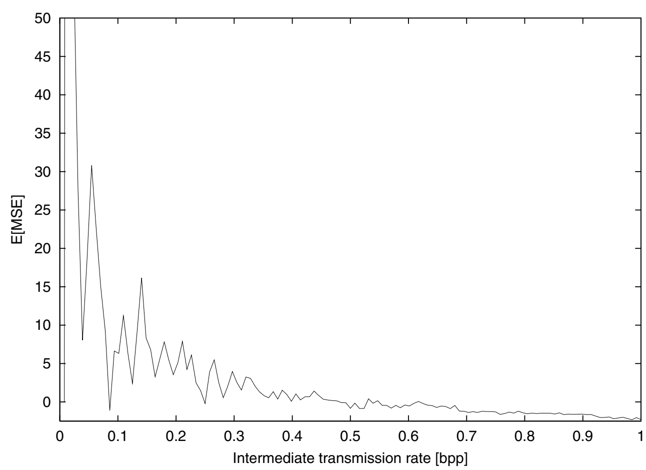

Fig. 5. Difference in expected MSE between CDO and PLS for JPEG 2000. The target transmission rate is 1 bpp.

TABLE III

EXPECTED MSE AT INTERMEDIATE TRANSMISSION RATES 0.25, 0.5, 0.75, AND 1 BPP FOR A CONSTRAINED DISTORTION-OPTIMAL SOLUTION (CDO), A PROGRESSIVE LOCAL SEARCH SOLUTION OPTIMIZED FOR ALL POSSIBLE INTERMEDIATE RATES (PLS), AND A PROGRESSIVE LOCAL SEARCH SOLUTION OPTIMIZED FOR THE ABOVE FOUR RATES (WPLS). THE TARGET TRANSMISSION RATE IS 1.0 BPP. THE SOURCE CODER IS SPIHT.

Interm. rate (bpp)	CDO E[MSE]	PLS E[MSE]	WPLS E[MSE]
0.25	86.6	81.55	72.36
0.5	37.34	37.13	36.65
0.75	25.64	26.39	24.52
1.0	18.03	19.56	18.36

ter performance than the solution that maximizes the average of the expected number of correctly received source bits. Compared to a solution that finds the optimal protection at the target transmission rate, the progressive local search solution had a slightly worse performance at transmission rates close to the target rate and a better performance at most of the intermediate rates.

Because Proposition 1 also holds for nondecreasing concave functions, the same methods discussed in the paper can be used to maximize the average expected PSNR.

We gave results for a binary symmetric channel. However the algorithm applies also to other wireless channel models, including the Gilbert-Elliott channel and the flat fading Rayleigh channel, which is a good model in wireless mobile communication. Future work will consider extensions to packet erasure channels.

Finally note that the same approach allows efficient progressive transmission of video sequences by using an embedded wavelet video bitstream [11] as a source code.

Acknowledgments. Vladimir Stanković thanks the Graduiertenkolleg "Wissensrepräsentation" (knowledge representation) of the Deutsche Forschungsgesellschaft (DFG) for funding. Youssef Charfi is supported by the German Academic Exchange Service (DAAD).

REFERENCES

[1] A. Said and W. A. Pearlman, *A new fast and efficient image codec based on set partitioning in hierarchical trees,* IEEE Trans. Circuits and Systems for Video Technology, vol. 6, pp. 243–250, June 1996.

[2] D.S. Taubman and M. Marcellin, *Jpeg2000: Image Compression Fundamentals, Standard, and Practice* Kluwer, 2001.

[3] P.G. Sherwood and K. Zeger, *Progressive image coding for noisy channels*, IEEE Signal Proc. Letters, vol. 4, no. 7, pp. 191–198, July 1997.

[4] P.G. Sherwood, X. Tian, and K. Zeger, *Channel code blocklength and rate optimization for progressive image transmission,* Proc. WCNC IEEE Wireless Communications and Networking Conference, pp. 978–982, 1999.

[5] V. Chande and N. Farvardin, *Progressive transmission of images over memoryless noisy channels*, IEEE Journal on Sel. Areas in Communications, Special Issue on Error Resilient Coding, vol. 18, no. 6, pp. 850–861, Jun 2000.

[6] V. Stanković, R. Hamzaoui, *Progressive optimal error protection of embedded codes*, Proc. of 9th Telecommunications Forum Telfor 2001, pp. 310–313, Belgrade, Nov. 2001.

[7] R. Hamzaoui, V. Stanković, and Z. Xiong. *Rate-based versus distortion-based optimal joint source-channel coding*, Proc. DCC'02 Data Compression Conference, J. A. Storer, M. Cohn (eds.), IEEE Computer Society Press, pp. 63–72, Snowbird, Utah, April 2002.

[8] V. Stanković, R. Hamzaoui, and D. Saupe, *Fast algorithm for optimal error protection of embedded wavelet codes*, Proc. MMSP-01 IEEE Workshop on Multimedia Signal Processing, pp. 593–598, Cannes, Oct. 2001.

[9] D.N. Rowitch and L.B. Milstein, *Rate compatible punctured turbo (RCPT) codes in a hybrid FEC/ARQ System*, Proc. Globecom'97, Phoenix, AZ, vol. 4., Nov. 1997, pp. 55–59.

[10] Y. Charfi, R. Hamzaoui, and D. Saupe, *Model-based real-time progressive transmission of images over noisy channels,* Proc. WCNC IEEE Wireless Communications and Networking Conference, New Orleans, March 2003.

[11] B. Kim, Z. Xiong, and W.A. Pearlman, *Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT),* IEEE Trans. Circuits Syst. Video Tech., vol. 10, number 8, pp. 692–695, Dec. 2000.