

Fast Forward Error Protection of Packetized Multimedia Bitstreams for Transmission over Varying Channels

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

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

Abstract— We propose a real-time optimization algorithm that selects an appropriate channel code for hybrid systems that combine packetization of an embedded wavelet bitstream into independently decodable packets and forward error correction using a family of channel codes with error detection and error correction capability. Such systems are very powerful for the transmission of audio, images, and video over fading and erasure channels with varying statistics. We also give an implementation that uses an optimal packetization technique and a concatenated cyclic redundancy check/rate-compatible punctured convolutional coder. Experimental results show that the peak signal-to-noise ratio of the average mean square error of our system is up to 1.74 dB higher than that of the previous best hybrid system for a Rayleigh fading channel and a transmission rate of 0.25 bits per pixel. Finally, we compare the hybrid approach to a state-of-the-art approach that uses a product code to protect the information bitstream.

I. INTRODUCTION

The embedded wavelet coders [1], [2], [3] have become popular in multimedia compression because of their competitive rate-distortion performance, low complexity, scalability, and progressive ability. However, these coders are very sensitive to channel noise because a single transmission failure (e.g., a bit error or a packet loss) can lead to an irreversible loss of synchronization between the encoder and the decoder. In one-way communication systems, two different approaches can be used to deal with this problem. The first one is based on forward error correction (FEC). A successful packet-based FEC technique for memoryless noisy channels was introduced by Sherwood and Zeger [4] who suggested to use a concatenation of an outer cyclic redundancy check (CRC) code for error detection and an inner rate-compatible punctured convolutional (RCPC) code for error correction. The system avoids error propagation by stopping the decoding when the first error in a packet is detected. Unfortunately, the performance of the system is not satisfactory for fading and erasure channels. The most powerful transmission systems over such channels are due to Sherwood and Zeger [5] and Sachs, Anand, and Ramchandran [6]. Both systems protect the embedded bitstream with a product channel code whose row code is a concatenated CRC/RCPC code, and column codes are systematic Reed-Solomon codes.

The second approach does not use FEC; it consists of reorganizing the output bitstream of the embedded wavelet coder

into packets that correspond to independent sets of wavelet coefficients [7], [8], [9]. Different strategies can be used to reconstruct wavelet coefficients from a lost packet. In [8], for example, they were recovered by interpolation if they belong to the low-low frequency band and set to zero otherwise. Since the packets are independently decodable, error propagation is limited to the packets where errors occur.

One important limitation of all above systems is that they are not efficient when the channel conditions are unknown or variable, which is the typical situation for a real-world channel. To overcome this problem, Cosman, Rogers, Sherwood, and Zeger [10] proposed a hybrid coder which combines the packetization technique of [8] with the FEC method of [4]. In this way, the output bitstream of an embedded wavelet coder is packetized into independently decodable packets, which are then protected by a concatenated CRC/RCPC coder. For pure Rayleigh flat-fading channels and those including packet erasure, the performance of the scheme outperforms that of [8] and [4] for channels with varying statistical conditions.

We propose two major improvements to the hybrid coder of [10]. In contrast to the approach of [10], which uses extensive simulations to select the RCPC code rate (the authors report 1000 trials per candidate code rate), we provide a real-time algorithm to find an appropriate RCPC code rate. Our algorithm finds an optimal code rate if the wavelet coefficients of the lost packets are quantized to zero. Moreover, experimental results show that the selected code rate is close to optimal if a more sophisticated interpolation technique is used. Second, we show that the performance of the original hybrid coder can be significantly improved by replacing the packetized zerotree wavelet coder of [8] with the optimal packetization scheme of [9].

The rest of the paper is organized as follows. In Section II, we present our fast algorithm for code rate selection. In Section III, we introduce our system and compare its performance to that of [10] and [6].

II. FAST ALGORITHM FOR CODE SELECTION

We consider a source-channel coding system that organizes the output bitstream of an embedded wavelet coder into independent blocks of the same size and protects all blocks with the same channel code taken from a family of channel codes with

error detection and error protection capability. The length of the channel codeword is fixed a priori and is equal to the size of the transmitted packet. Thus, the length of the information block is determined by the selected channel code rate.

Let $\mathcal{R} = \{r_1, \dots, r_m\}$ be the set of channel code rates with $r_1 < \dots < r_m$. Suppose that N packets of length L each are successively sent over an unreliable channel where both packet loss and bit errors can occur. If the receiver detects an error in a packet, then this packet is considered to be lost (we suppose that all errors can be detected). Coefficients from the lost packets are either determined by interpolation, using the information from the correctly recovered packets, or set to zero.

Let $Q = \{1, 2, \dots, N\}$ be the set of packet indices (packet 1 is sent first, packet 2 is sent next, etc). Consider all possible subsets of n elements of Q and let Q_n^j , $n = 1, \dots, N$, $j = 1, \dots, \binom{N}{n}$, denote the j th such subset. For example, if $Q = \{1, 2, 3, 4\}$, then $Q_1^1 = \{1\}$, $Q_2^1 = \{2\}$, $Q_3^1 = \{3\}$, $Q_4^1 = \{4\}$, $Q_2^2 = \{1, 2\}$, $Q_2^3 = \{1, 3\}$, $Q_2^4 = \{1, 4\}$, $Q_2^5 = \{2, 3\}$, $Q_2^6 = \{2, 4\}$, $Q_2^7 = \{3, 4\}$, $Q_3^1 = \{1, 2, 3\}$, $Q_3^2 = \{1, 2, 4\}$, $Q_3^3 = \{1, 3, 4\}$, $Q_3^4 = \{2, 3, 4\}$, and $Q_4^1 = \{1, 2, 3, 4\}$.

Suppose that the packets are protected with channel code rate $r_k \in \mathcal{R}$. Let $P^{(r_k)}(Q_n^j)$ denote the probability that all packets in Q_n^j are lost, while all other packets are correctly decoded, and let $d^{(r_k)}(Q_n^j)$ denote the corresponding distortion. Then the expected distortion is

$$E_N[d](r_k) = P_0^{(r_k)} d_0^{(r_k)} + \sum_{n=1}^N \sum_{j=1}^{\binom{N}{n}} P^{(r_k)}(Q_n^j) d^{(r_k)}(Q_n^j), \quad (1)$$

where $P_0^{(r_k)}$ and $d_0^{(r_k)}$ are the probability that no packets are lost and the corresponding distortion, respectively.

Finding a channel code rate r_k that minimizes (1) cannot be done by exhaustive search for large N . However, by exploiting the fact that each packet contains independent bitstreams, the problem can be simplified as follows. Let $n \leq N$ and $1 \leq l_1 < \dots < l_n \leq N$. For $Q_n^j = \{l_1, \dots, l_n\}$, assume that

$$d^{(r_k)}(Q_n^j) = d^{(r_k)}(Q_1^{l_1}) + \dots + d^{(r_k)}(Q_1^{l_n}) - (n-1)d_0^{(r_k)}, \quad (2)$$

which is satisfied if all lost wavelet coefficients are set to zero and only an approximation if a sophisticated interpolation technique is used.

For $l, n \in \{1, \dots, N\}$, let $P^{(r_k)}(l; n)$ denote the probability that packet l is lost and the total number of lost packets is n . Since all packets are of equal length, for any positive integers $i, j \leq N$ we have $P^{(r_k)}(i; n) = P^{(r_k)}(j; n) = P^{(r_k)}(1; n)$. Using this and (2), we can rewrite the right-hand side of (1) as

$$P_0^{(r_k)} d_0^{(r_k)} + \sum_{n=1}^N P^{(r_k)}(1; n) (\mathbf{d}^{(r_k)} + N d_0^{(r_k)}) - (n-1) P_n^{(r_k)} d_0^{(r_k)} \quad (3)$$

where $P_n^{(r_k)}$ is the probability that n packets of N are lost and $\mathbf{d}^{(r_k)} = \sum_{j=1}^N (d^{(r_k)}(Q_1^j) - d_0^{(r_k)})$. But $P^{(r_k)}(1; n) = \frac{n}{N} P_n^{(r_k)}$. Thus (3) reduces to

$$d_0^{(r_k)} + \mathbf{d}^{(r_k)} \sum_{n=1}^N \frac{n}{N} P_n^{(r_k)}. \quad (4)$$

The distortion $d_0^{(r_k)}$ and the sum $\mathbf{d}^{(r_k)}$ are channel independent. On the other hand, the probabilities $P_n^{(r_k)}$ depend only on the channel statistics and are image independent. When these values are known, minimizing (4) can be done in real-time. The optimization can additionally be accelerated by discarding a priori many code rates. Indeed, for each $r_k \in \mathcal{R}$, we have $E_N[d](r_k) \geq d_0^{(r_k)}$. Thus, all code rates r_k such that $d_0^{(r_k)}$ is greater than the temporary minimal expected distortion can be discarded. Note also that for $r_i, r_j \in \mathcal{R}$, $r_i > r_j$ gives $d_0^{(r_i)} < d_0^{(r_j)}$. We now give the pseudocode of the algorithm.

Algorithm 1: An optimal code rate r_{\min} is obtained as follows.

1. Set $n = m$, $\min = E_N[d](r_n)$, and $r_{\min} = r_n$.
2. Set $n = n - 1$. If $n = 0$, stop. If $d_0^{(r_n)} > \min$, stop.
3. If $\min > E_N[d](r_n)$, then set $\min = E_N[d](r_n)$ and $r_{\min} = r_n$.
4. Go to Step 2.

Experimental results in Section III show that the assumption (2) is acceptable for the interpolation technique of [9].

III. RESULTS

We propose to modify the system of [10] as follows:

- we replace the packetization method of [8] by the optimal packetization method of [9];
- we select the RCPC code rate with Algorithm 1 instead of using extensive simulations;
- we fix the transmission packet size to 48 bytes. Thus, the size is independent of the selected code rate. This is more suitable in practice than the approach of [10] where the information block size is fixed to 48 bytes while each code rate selects a different transmission packet size.

We show in this section that our system has a better performance than the original hybrid system of [10], and that it is more robust against varying channel conditions than the system of [6]. Robustness over varying channels is important in wireless communication where channel statistics rapidly change in time; thus, very often, a system designed for one channel state is used at another. We also show that Algorithm 1 allows real-time selection of an appropriate RCPC code rate.

We first compare the performance of our system to that of [10] for a flat-fading Rayleigh channel. As in [10], we used a 16-bit CRC with generator polynomial 0x15935. The generator polynomials of the RCPC were (0117, 0127, 0155, 0171), the mother code rate was 1/4, and the puncturing rate was 8 [11]. Thus, the set of RCPC code rates was $\{8/9, 8/10, \dots, 8/32\}$. The decoding of the RCPC code was done with a list Viterbi algorithm where the maximum number of candidate paths was 100. The two systems were optimized for a target transmission rate of 0.25 bits per pixel (bpp) and a Rayleigh fading channel with a signal-to-noise ratio (SNR) of 13 dB and a normalized Doppler spread $f_D = 10^{-4}$. The optimization was done under the constraint that at least 95 % of the decoded images have a mean-square error (MSE) lower than 818.6 (19 dB in peak signal-to-noise ratio (PSNR)), which, as specified in [10], is the lowest bound for a successful recognition of an image. Table I shows the average MSE at various channel conditions for the

$f_D(\times 10^{-6})$	System	Average received SNR (dB)			
		10	12	15	20
10	[10]	160.5 (26.08)	136.6 (26.78)	115.7 (27.50)	96.9 (28.27)
	Proposed	127.4 (27.08)	100.4 (28.11)	78.8 (29.17)	64.3 (30.05)
50	[10]	181.9 (25.53)	147.5 (26.44)	112.7 (27.61)	91.4 (28.52)
	Proposed	122.8 (27.24)	98.9 (28.18)	78.6 (29.18)	64.4 (30.04)
100	[10]	163.3 (26.00)	122.0 (27.27)	99.6 (28.15)	89.0 (28.64)
	Proposed	124.8 (27.17)	100.9 (28.09)	80.0 (29.10)	65.4 (29.97)
200	[10]	126.4 (27.11)	103.8 (27.97)	92.3 (28.48)	88.2 (28.68)
	Proposed	129.9 (27.00)	105.3 (27.91)	83.6 (28.91)	67.1 (29.86)

TABLE I

AVERAGE MSE OF THE LENA IMAGE FOR OUR SYSTEM AND THE SYSTEM OF [10] AT TRANSMISSION RATE 0.25 BPP. THE OPTIMIZATION WAS DONE FOR A RAYLEIGH CHANNEL WITH $f_D = 100 \times 10^{-6}$ AND SNR = 13 DB. THE PSNR OF THE AVERAGE MSE IS SHOWN IN THE BRACKETS.

Packet loss rate	Burst length	[10]	Proposed
0.01	1	137.6 (26.74)	108.05 (27.79)
0.1	1	206.9 (24.97)	176.59 (25.66)
0.01	10	136.6 (26.78)	108.22 (27.79)
0.1	10	200.5 (25.11)	176.82 (25.66)

TABLE II

AVERAGE MSE OF THE LENA IMAGE AT TRANSMISSION RATE 0.25 BPP. THE OPTIMIZATION WAS DONE FOR A PURE RAYLEIGH CHANNEL WITH $f_D = 10^{-4}$ AND SNR = 13 DB. RESULTS ARE FOR THE COMBINATION OF A RAYLEIGH CHANNEL (SNR = 12 DB, $f_D = 10^{-4}$) AND A PACKET ERASURE CHANNEL WITH VARIOUS LOSS RATES AND BURST LENGTHS.

SPIHT bitstream of the 8 bpp 512×512 Lenna image. Our results were averaged over 5000 simulations, whereas the results of [10] were reported as the average of 1000 to 3000 experiments. The Rayleigh channel was simulated with Jakes' [12] channel model. Table II shows results for transmission over a combination of a Rayleigh fading channel and a packet erasure channel. The performance of our system was almost always better than that of [10]. The PSNR gain was up to 1.74 dB over the pure Rayleigh channel and up to 1.05 dB over the combination of the Rayleigh and erasure channels.

We now compare our system to the product code approach [5], [6]. We used the system of [6] in our experiments because it provides a better performance than that of [5]. For a fair comparison between our system and that of [6], the same CRC/RCPC coder specified above was used. Note, however, that the product code approach was advantaged over the hybrid approach because it included adaptive arithmetic coding of the source bitstream, which is not possible with packetization since packet loss destroys context adaptation.

When the optimization was done for the exact channel conditions, the system of [6] had a significantly better performance. For example, for the Rayleigh fading channel with $f_D = 10^{-4}$ and SNR = 13 dB, the PSNR of the average MSE of our system was 28.48 dB while that of [6] was 31.54 dB. However, the situation changed when the channel conditions varied. Table III shows results for the Lenna image in Rayleigh channels with SNR = 10 dB and four different values of f_D .

Figure 2 shows results when f_D was fixed to 10^{-5} , and the SNR varied in the range 10 to 20 dB. Whereas our system was robust against changing channel conditions, the system of [6]

System	Normalized Doppler spread $f_D (\times 10^{-6})$			
	10	50	100	200
Proposed	27.08	27.24	27.17	27.00
[6]	21.45	24.04	26.99	28.44

TABLE III

PSNR IN DB OF THE AVERAGE MSE FOR THE LENA IMAGE. BOTH SYSTEMS WERE OPTIMIZED FOR A RAYLEIGH CHANNEL WITH $f_D = 10^{-4}$ AND SNR = 13 DB AND TESTED AT SNR = 10 DB AND FOUR DIFFERENT f_{DS} . THE TRANSMISSION RATE IS 0.25 BPP.

collapsed when long bursts (low f_D and SNR) were present. A similar behavior was observed for other channel parameters. As noted in [10], in addition to the average decoded distortion, one should also consider the cumulative distribution function of the decoded distortion to measure the performance of a source channel coding system in a noisy channel. Figure 1 indicates that for varying channels our system is less likely to generate poor reconstructions than the system of [6]. However, since the performance of the SPIHT bitstream decreases after packetization, the system of [6] had a higher probability of generating the lowest MSEs. For example, the lowest MSE produced by our system was 60, whereas the system of [6] yielded an MSE of 48 in 60 % of the cases.

Figure 3 shows visual results. They confirm the observation made in [10] that reconstruction errors in the hybrid scheme (packetization and CRC/RCPC) are localized in space, allowing a better subjective quality than that of the pure CRC/RCPC coder where the error is distributed over the whole image.

In Table IV, we compare for our system the solutions computed with Algorithm 1 to those obtained by simulations. Optimization by simulations consisted of selecting the RCPC code rate that gave the lowest average distortion in 5000 independent experiments. The results are shown for a pure Rayleigh channel and for a combination Rayleigh/packet erasure channel. The test images were the 512×512 Lenna and Peppers. The transmission rate was 0.25 bpp. Whereas the optimization by simulations is very time-consuming and thus, impractical for online multimedia communications, Algorithm 1 required only 0.02 milliseconds on a PC having an AMD Athlon (TM) XP 1600+ 1400 MHz processor with a main memory size of 1 Gbyte. Moreover, the experiment shows that Algorithm 1 pro-

Packet loss rate	Image	Method	Average received SNR (dB)			
			10	12	15	20
0	Lenna	Algorithm 1	118.78 (8/14)	99.71 (8/13)	79.11 (8/11)	56.69 (8/9)
		Simulations	118.78 (8/14)	99.71 (8/13)	79.11 (8/11)	56.69 (8/9)
0	Peppers	Algorithm 1	141.32 (8/14)	116.19 (8/13)	91.39 (8/12)	66.03 (8/9)
		Simulations	141.32 (8/14)	116.19 (8/13)	91.39 (8/12)	66.03 (8/9)
0.1	Lenna	Algorithm 1	194.55 (8/14)	175.60 (8/13)	156.49 (8/11)	134.59 (8/9)
		Simulations	194.37 (8/15)	175.43 (8/14)	156.49 (8/11)	134.59 (8/9)
0.1	Peppers	Algorithm 1	243.04 (8/14)	217.91 (8/13)	193.52 (8/12)	168.58 (8/9)
		Simulations	243.04 (8/14)	217.91 (8/13)	193.52 (8/12)	168.31 (8/10)

TABLE IV

AVERAGE MSE AT 0.25 BPP FOR A RAYLEIGH CHANNEL WITH $f_D = 10^{-4}$. THE SELECTED CODE RATES ARE SHOWN IN THE BRACKETS.

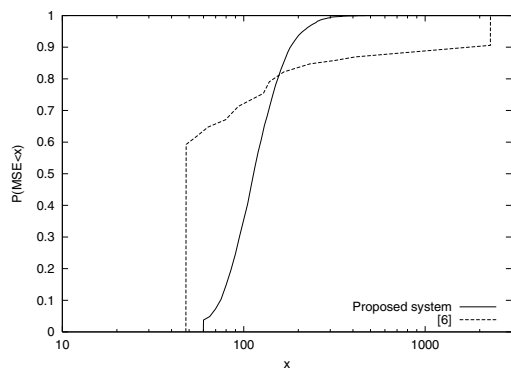


Fig. 1. Cumulative distributions of the distortion for the Lenna image. The code rates were optimized for a Rayleigh channel ($f_D = 10^{-4}$, SNR = 13 dB). Results are shown for the combination of a packet erasure channel with packet error rate 0.01 and a Rayleigh channel with SNR = 10 dB and $f_D = 5 \times 10^{-5}$. The transmission rate is 0.25 bpp.

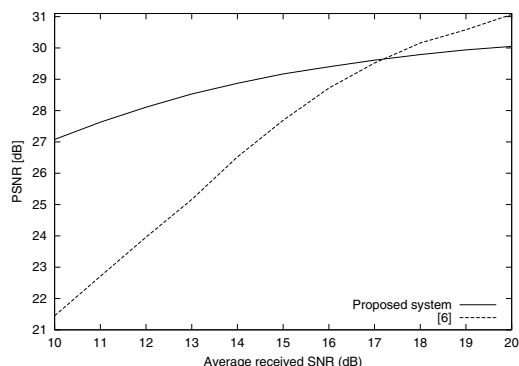


Fig. 2. PSNR in dB of the average MSE vs. average received SNR for the Lenna image. Both systems were optimized for a Rayleigh channel with $f_D = 10^{-4}$ and SNR = 13 dB, and tested over Rayleigh channels with $f_D = 10^{-5}$ and different SNRs. The transmission rate is 0.25 bpp.

vides either optimal or near-optimal solutions.

IV. CONCLUSION

The major contribution of our paper is a real-time algorithm for code rate selection in systems that combine packetization of an embedded wavelet bitstream and forward error correction with a family of channel codes having error detection and error correction capability. Such systems were introduced by Cosman *et al.* [10] who showed their efficiency for protecting images in channels with varying conditions. Our algorithm can be used with any packetization method, including those of

[9], [7], and [8]. It finds an optimal code rate if the receiver approximates the wavelet coefficients of the lost packets by zeroes. Experimental results showed that the selected code rate is near-optimal if the interpolation method of [9] is used.

Another contribution was to significantly improve the performance of the scheme of [10] by using the optimal packetization method of [9] instead of the packetization technique of [8].

No comparison between the product code approach and the hybrid approach was available. This motivated our last experiment, where we compared our system to the state-of-the-art product code scheme of [6]. Our system had a worse average end-to-end performance in a fading channel with known statistics, but it was much more robust to varying channel conditions. A promising topic for future research may be the design of a system that uses both the optimal packetization technique of [9] and a product code.

We provided experimental results for images; note, however, that our system can be used to protect embedded video codes [3] as well.

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 Syst. Video Tech.*, vol. 6, pp. 243–250, June 1996.
- [2] D.S. Taubman and M. Marcellin, *Jpeg2000: Image Compression Fundamentals, Standards, and Practice*, Kluwer, 2001.
- [3] 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, no. 8, pp. 692–695, Dec. 2000.
- [4] P.G. Sherwood and K. Zeger, *Progressive image coding for noisy channels*, *IEEE Sig. Proc. Letters*, vol. 4, no. 7, pp. 191–198, July 1997.
- [5] P.G. Sherwood and K. Zeger, *Error protection for progressive image transmission over memoryless and fading channels*, *IEEE Trans. Comm.*, vol. 46, no. 12, pp. 1555–1559, Dec 1998.
- [6] D.G. Sachs, R. Anand, and K. Ramchandran, *Wireless image transmission using multiple-description based concatenated codes*, *Proc. SPIE 2000*, vol. 3974, pp. 300–311, San Jose, CA, Jan. 2000.
- [7] C. Creusere, *A new method of robust image compression based on the embedded zerotree wavelet algorithm*, *IEEE Trans. Image Proc.*, vol. 6, no. 10, pp. 1436–1442, Oct. 1997.
- [8] J. Rogers and P. Cosman, *Wavelet zerotree image compression with packetization*, *IEEE Sig. Proc. Letters*, vol. 5, no. 5, pp. 105–107, May 1998.
- [9] X. Wu, S. Cheng, and Z. Xiong, *On packetization of embedded multimedia bitstreams*, *IEEE Trans. Multimedia: Special Issue on Multimedia over IP*, vol. 3, pp. 132–140, March 2001.
- [10] P. Cosman, J. Rogers, P.G. Sherwood, and K. Zeger, *Combined forward error control and packetized zerotree wavelet encoding for transmission of images over varying channels*, *IEEE Trans. Image Proc.*, vol. 9, pp. 982–993, June 2000.
- [11] P. Frenger, P. Orten, T. Ottosson, and A. Svensson, *Multi-rate Convolutional Codes*, Technical report R021/1998, Göteborg, 1998.
- [12] W.C. Jakes, *Microwave Mobile Communications*, Wiley, 1974.



27.15 dB



26.99 dB



25.65 dB



25.27 dB



25.68 dB



22.12 dB

Fig. 3. Median reconstruction quality of our system (left) and the system of [6] (right) at transmission rate 0.25 bpp. Both systems were optimized for a Rayleigh channel with $f_D = 10^{-4}$ and SNR = 13 dB and tested at different channel conditions. First row: $f_D = 10^{-4}$, SNR = 10 dB, no packet erasure. Second row: $f_D = 10^{-4}$, SNR = 12 dB, the packet erasure rate is 0.1 and the burst length 1. Third row: as in the second row but the burst length is 10.