Joint Source-Channel Image Coding for a Power Constrained Noisy Channel *

Marc P.C. Fossorier   Zixiang Xiong   Kenneth Zeger

Dept. of Electrical Engineering
University of Hawaii
Honolulu, HI 96822

Dept. of Electrical and Computer Eng.
University of California, San Diego
La Jolla, CA 92093-0407

Abstract

We study joint source-channel coding for a power constrained Gaussian channel and its application to progressive image compression. For a given power constrained, we consider the optimum allocation of energy per bit for a BPSK transmitter and the best choice of channel code rate, when the performance is measured by end-to-end average quantizer distortion. Choosing the average energy per transmitted bit in conjunction with both the source rate and the channel code rate provides an additional degree of freedom with respect to previously proposed schemes, and therefore can achieve higher overall PSNRs for images.

1 Introduction

For a source (such as an image) with distortion-rate function $D(\cdot)$ and a binary symmetric channel with capacity C, Shannon's "separation principle" ensures that transmission of the source at a rate of R bits/sample over the noisy channel can be achieved with a distortion arbitrarily close to $D(RC)$, by choosing independently the source and channel coders. However, this theoretical result assumes unboundedly long block lengths and unrestricted computational complexity. In practice, delay and complexity constraints motivate the search for source and channel codes which efficiently trade off the available transmission rate between source coding and channel coding.

The "cost" of using a discrete channel is generally described in terms of the transmission rate, measured in channel uses (i.e. bits sent) per source symbol. In contrast, the "cost" of using a power constrained non-discrete channel (with unlimited bandwidth), such as the additive white Gaussian noise (AWGN) channel, is described by the average energy transmitted per source symbol. That is, the number of signal constellation points transmitted per source symbol is a design parameter that can be chosen to optimize the end-to-end mean squared quantization error of the system. If one chooses a higher transmission rate, then there is less energy transmitted per bit, and thus a larger bit error probability results per transmitted bit. There is thus a trade-off between the transmission rate, the source coding rate, and the channel code rate. In this paper, we describe an optimization over this triple of parameters. In particular, we demonstrate this optimization for the case when the source coder is a progressive zerotree wavelet-based coder (SPIHT), the channel coder chosen from a family of BCH block codes, and the modulator used is binary phase shift keying (BPSK).

Under high resolution assumptions, an optimum trade-off between fixed-delay source coding and block channel coding was derived in [1] for the binary symmetric channel (BSC), and in [2] for the Gaussian channel. The derivations in [1] and [2] assumed the usage of error correcting codes achieving exponentially small probability of error in terms of block length. Such codes are known to exist but it is not known how to find and use them.

In [3], an effective coding scheme was presented for image transmission over a BSC, by combining Said and Pearlman's improvement [5] of Shapiro's wavelet source coding algorithm [4], and Hagenauer's rate compatible punctured convolutional (RCPC) channel codes [6]. The results in [3] demonstrate that by choosing the best available source and channel coders independently, and wisely selecting the corresponding source and channel code rates, very good results can still be obtained.

In this paper, we generalize the scheme of [3] from a discrete channel model to an analog channel. For a given power constrained Gaussian communication channel, we consider the optimum allocation of en-

*This work was supported in part by the National Science Foundation.

ergy per bit for a binary phase shift keying (BPSK) transmitter when the performance is measured by end-to-end average quantizer distortion. Choosing the average energy per transmitted bit in conjunction with both the source rate and the channel code rate provides one additional degree of freedom with respect to previously proposed schemes for discrete value channels, and therefore achieves higher overall peak signal to noise ratio (PSNR) values for images. As an example, for transmission of the 512 x 512 image "Lena" over a BSC with crossover probability $p = 10^{-3}$, a coding gain of about 1 dB over the results of [3] is achieved by properly choosing the averaging energy per transmitted bit.

2 Source and Channel Coding Trade-off with a Power Constraint

Suppose that the quality of a sampled and encoded analog source is characterized by its distortion $\mathcal{D}(r_s)$ as a function of its rate r_s (measured in bits per source sample). The quantity $\mathcal{D}(r_s)$ typically measures the mean-square quantization error of a decompressed image in terms of the number of bits per pixel (bpp) present in the compressed version of the image. Suppose a channel code with rate r_c acts on the output bits from the source encoder. The resulting bit stream is transmitted across an AWGN channel with zero-mean and variance $N_0/2$ using a BPSK modulator whose decoding is performed by a hard-limiter on the received sampled values. Suppose the BPSK modulator emits a sequence of constellation points (analog values) $x_1, x_2, x_3, \cdots$ with $x_i \in \{-\sqrt{P}, \sqrt{P}\}$ where P is a fixed power constraint (measured in units of energy per source sample). Consequently, if R constellation signals per source sample are transmitted over the channel, the average energy E_s per transmitted signal satisfies

$$R \cdot E_s = P.$$

The number of bits per source sample available for source coding is $r_s = R \cdot r_c$ and the probability of error for a transmitted channel bit is $p(E_s) = Q(\sqrt{2E_s/N_0})$, where

$$Q(x) = (2\pi)^{-1/2} \int_x^\infty e^{-u^2/2} du.$$

$p(E_s)$ is the crossover probability of the resulting discrete BSC. The capacity of this BSC is

$$C(E_s) = 1 - h(p(E_s)), \qquad (1)$$

where $h(\cdot)$ is the binary entropy function, defined by

$$h(x) = -x \log_2(x) - (1-x) \log_2(1-x).$$

Shannon's channel coding theorem shows that for a fixed E_s, if $r_c < C$ then r_s bits per source sample can be transmitted with arbitrarily small probability of error, and Shannon's separation principle shows that the distortion $\mathcal{D}(r_s)$, corresponding to rate r_s, can be achieved. Shannon's theorem predicts the theoretically achievable reliable transmission rate, although it assumes unboundedly long block lengths.

In [3], the following related problem for image transmission is considered: *Given a BSC with crossover probability p and a fixed transmission rate R_0 over this channel, what is the best possible distortion achievable at the receiver with channel coding?* If we assume transmission over an AWGN channel, then the value N_0 is implicitly fixed. It follows that the energy E_s per BPSK transmitted channel signal takes the fixed value $E_b = (N_0/2)(Q^{-1}(p))^2$, where $Q^{-1}(\cdot)$ represents the inverse function of $Q(\cdot)$. Also, the power constraint is implicitly defined by $P = R_0 E_b$.

In contrast, in this paper, we fix the power constraint to P and allow E_s to vary, so that for

$$\alpha = E_s/E_b,$$

the transmission rate becomes $R = P/E_s = R_0/\alpha$. Consequently, for each channel code considered, E_s is optimized instead of being taken as a given. The model of [3] corresponds to the case $\alpha = 1$. A value $\alpha < 1$ can be viewed as a degradation of the channel to allow the use of a more powerful channel code, and conversely for $\alpha > 1$. Note finally that with this new model, the time required to transmit the source output considered remains the same as in [3].

3 Trade-off Optimization for Images

In this section, we consider families of implementable codes (in contrast to Shannon's random codes), chosen based on specific design criteria. For the family of codes considered, we then seek to solve the problem: *For a fixed AWGN channel (fixed N_0) and a given power constraint (given P), jointly determine the code of rate r_c and the average energy E_s per transmitted BPSK signal that minimize the average mean-squared error (MSE) of a transmitted image.*

In the following, we study this practical problem using the progressive image coders of [4, 5]. The embedded nature of these schemes particularly fits our problem as whenever the channel decoder fails to correct an error, all the previous decoded blocks for the transmitted image considered can still be used to obtain a different resolution version of this image. Also, once a channel decoding error has occurred, source decoding stops. For a given AWGN channel with average power constraint P, we determine the best trade-off between source and channel coding to minimize the average MSE of a transmitted image using a block channel code of fixed length n. We assume Said and Pearlman's wavelet source coding scheme is used in conjunction with an (n, k, d) BCH channel code of length n, dimension k, and minimum Hamming distance d [7]. We use BCH codes (instead of RCPC codes as in [3]) since they allow a tight error performance analysis by means of their known distance distribution. Also, efficient algebraic decoders for BCH codes have been devised, so that they are both of practical and theoretical interest. However, the following analysis remains valid for any linear block channel code for communication over a BSC.

For an image of size M pixels transmitted at a rate $R = P/E_s$ bpp, the number of blocks of k bits each that are encoded by the (n, k, d) code is

$$b(E_s) = \left\lceil \frac{Mr_s}{k} \right\rceil = \left\lceil \frac{MR}{n}. \right\rceil \quad (2)$$

If we set $t = \lfloor (d-1)/2 \rfloor$, the corresponding block error probability after decoding is

$$P_s \leq \sum_{i=t+1}^{n} \binom{n}{i} p(E_s)^i (1 - p(E_s))^{n-i}, \quad (3)$$

with equality if only the errors of weight strictly less than $t + 1$ are corrected (which is assumed in the following) [7]. Due to the embedded nature of the source coder, the average MSE (average over both source and channel statistics) can be expressed explicitly as a function of r_c and E_s as

$$\begin{aligned} \delta(r_c, E_s) &= \mathcal{D}(r_s)(1 - P_s)^{b(E_s)} \\ &+ \sum_{i=0}^{b(E_s)-1} \mathcal{D}\left(r_s \frac{i}{b(E_s)}\right)(1 - P_s)^i P_s, \end{aligned} \quad (4)$$

where $\mathcal{D}(r)$ represents the distortion of the image compressed at a rate of r bpp with no channel noise. In (4), $\mathcal{D}(r_s i/b(E_s))$ corresponds to the distortion achieved when the source decoder stops after i blocks due to a first channel decoder error at the $(i + 1)$-th block.

For a given channel code of rate r_c, we have $P_s \to 0$ as $\alpha \to \infty$ in which case $\lim_{\alpha\to\infty} \delta(r_c, E_s) = \lim_{\alpha\to\infty} \mathcal{D}(R_0 r_c/\alpha) = \mathcal{D}(0)$. Also, $P_s \to 1$ as $\alpha \to 0$, so that $\delta(r_c, E_s)$ approaches $\mathcal{D}(0)$ for fixed values of M and n. Hence, the minimum value of $\delta(r_c, E_s)$ is achieved for some finite nonzero choice of α.

For each (n, k, d) code of the family considered with $r_c = k/n$, the value E_s providing the minimum MSE value $\delta_{min}(r_c)$ in (4) is chosen. The corresponding source code rate is $r_s = Pr_c/E_s = R_0 r_c/\alpha$. Finally, the code of rate r_c with minimum MSE value δ_{min} is chosen, so that

$$\begin{aligned} \delta_{min} &= \min_{r_c}\{\min_{E_s}\{\delta(r_c, E_s)\}\}, \\ &= \min_{r_c}\{\delta_{min}(r_c)\}. \end{aligned} \quad (5)$$

In the approach of [3], the code chosen corresponds to the MSE

$$\delta_{min}(E_b) = \min_{r_c}\{\delta(r_c, E_b)\}. \quad (6)$$

4 Simulation results

Consider the transmission of the 512×512 $(= M)$ image "Lena" over a BSC with crossover probability $p(E_b)$ for the transmission rate R_0. Figure 1 depicts the PSNR values achieved by seven BCH codes of length $n = 127$ for $p(E_b) = 10^{-3}$ and $R_0 = 1$ bpp. We observe that the (127,71) BCH code allows us to achieve δ_{min} and a PSNR value of 40.38 dB by choosing $E_s = 0.54E_b$ and adjusting $r_s = (71/127) \cdot 0.54^{-1} = 1.035$ bpp. In comparison, the PSNR value corresponding to $R_0 = 1$ bpp is 40.40 dB in the noiseless channel case, indicating that very good performance has been achieved by the proposed scheme.

Although $r_c = 71/127$ and $\alpha = 0.54$ define the optimum choice for the family of codes represented in Figure 1, it is important to notice that in practice, choosing $E_s > 0.54E_b$ provides a choice that is more robust to variable channel conditions, due to the large slope of the curve on the left of the maximum. We finally observe that for this family of codes, our approach provides more than 1 dB gain over the case $E_s = E_b$ corresponding to [3], with the choice of a different channel code.

For the same family of length 127 codes, Figure 4 depicts the results for $p(E_b) = 10^{-2}$ and $R_0 = 0.25$ bpp. Again, the (127,71) BCH code provides the largest PSNR value of 32.04 dB for $\alpha = 0.85$. However, other codes of this family as well as the scheme of [3] for the same code allow us to achieve close PSNR values. Also, the highest PSNR achieved by this family of codes is more than 2 dB lower than that corresponding to $R_0 = 0.25$ bpp in the noiseless case, which suggests that for this relatively noisy BSC, a family of more powerful codes of longer lengths have to be considered to approach this value. Interestingly, due to the relatively flat shapes of the curves around their maxima, the optimum choice of α becomes much less subject to variable channel conditions than for the schemes of Figure 1. Table 1 summarizes the optimum choices of α for the family of length $n = 127$ codes, $p(E_b) = 10^{-2}$ and 10^{-3}, and $R_0 = 0.25$ and 1 bpp. The corresponding simulation results are depicted in Figures 1, 2, 3 and 4.

In all previous four cases considered, the highest PSNR values obtained remain lower than that corresponding to R_0 in the noiseless case. Note however that by allowing E_s to change, it is possible to achieve an even higher PSNR value than that corresponding to R_0 in the noiseless case. Such is the case for the schemes depicted in Figure 5 for $p(E_b) = 10^{-6}$ and $R_0 = 0.25$ bpp. For the (127,71) BCH code, choosing $\alpha = 0.23$ and $r_s = 0.608$ bpp allows us to achieve a PSNR value of 37.97 dB against 34.11 dB for $R_0 = 0.25$ bpp in the noiseless case. This gain is possible due to the good quality of the BSC considered.

5 Conclusion

These results can be extended to other types of channels or modulation forms. For example, for the same family of codes simulated in this paper, better results can be obtained by considering the AWGN channel and soft decision decoding, rather than the corresponding BSC and hard decision decoding, at the expense of a much larger decoding complexity for the channel decoder. However, for such schemes, (3) becomes a strict union upper bound, so that $\delta(r_c, E_s)$ can only be upper bounded by (4). As a result, the exact PSNR values can in general only be determined by simulations, especially for values $P_s \geq 10^{-4}$ for which the union bound remains quite loose. Note that the channel coding scheme of [3] based on the concatenation of CRC and RCPC codes for the BSC is subject to the same analytical problem due to the facts that the error capability of the concatenated code is unknown and the proposed decoding is suboptimum.

References

[1] B. Hochwald and K. Zeger, "Tradeoff Between Source and Channel Coding," *IEEE Trans. on Inform. Theory*, vol. 43, pp. 1412-1424, September 1997.

[2] B. Hochwald, "Tradeoff Between Source and Channel Coding on a Gaussian Channel," *IEEE Trans. on Inform. Theory*, 1998 (to appear).

[3] P. G. Sherwood and K. Zeger, "Progressive Image Coding for Noisy Channels," *IEEE Signal Processing Letters*, vol. 4, pp. 189-191, July 1997.

[4] J. M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients", *IEEE Trans. on Signal Processing*, Vol. 41, pp. 3445-3463, December 1993.

[5] A. Said and W. A. Pearlman, "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees", *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 6, pp. 243-250, June, 1996.

[6] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. on Communications*, vol. 36, pp. 389-400, April 1988.

[7] S. Lin and D. J. Costello Jr, *Error Control Coding Fundamentals and Applications*, New Jersey: Prentice-Hall, 1983.

	$R_0 = 0.25$	$R_0 = 1.0$
$p(E_b) = 10^{-3}$	34.38 dB (α= 0.5)	40.38 dB (α=0.54)
$p(E_b) = 10^{-2}$	32.04 dB (α= 0.85)	37.83 dB (α=0.93)

Table 1: Coding results for the 512×512 Lena image (in each case, the (127, 71) code performs the best among the class of length 127 BCH codes).

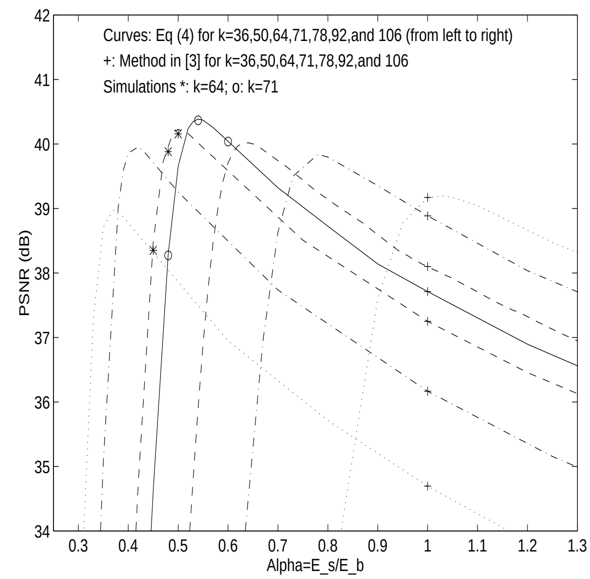

Figure 1: PSNR values for different BCH codes of length 127, for $p(E_b) = 10^{-3}$ and $R_0 = 1$ bpp.

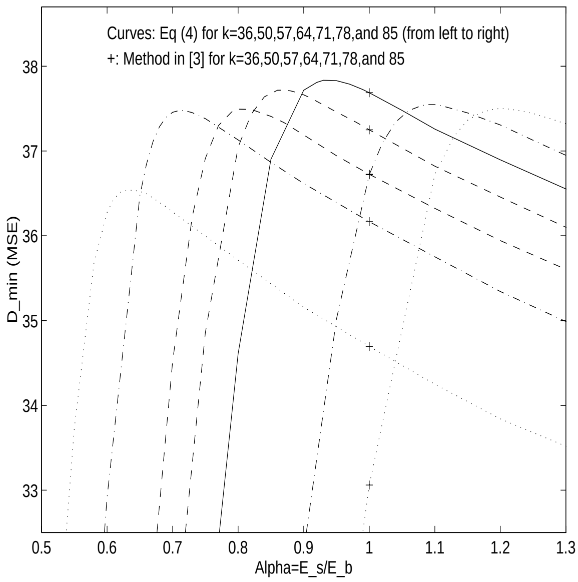

Figure 2: PSNR values for different BCH codes of length 127, for $p(E_b) = 10^{-2}$ and $R_0 = 1$ bpp.

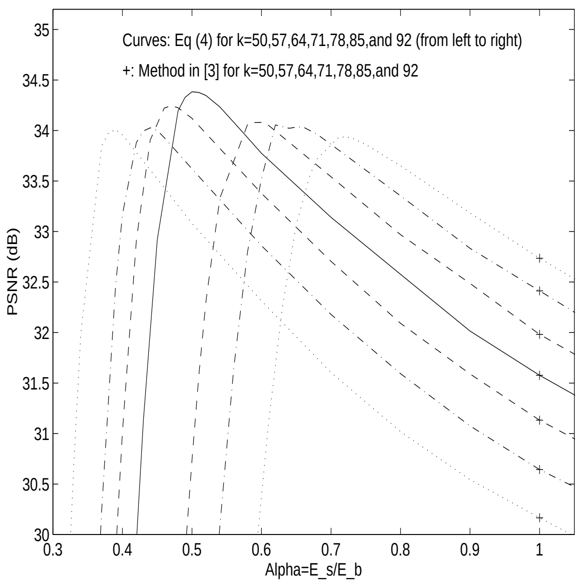

Figure 3: PSNR values for different BCH codes of length 127, for $p(E_b) = 10^{-3}$ and $R_0 = 0.25$ bpp.

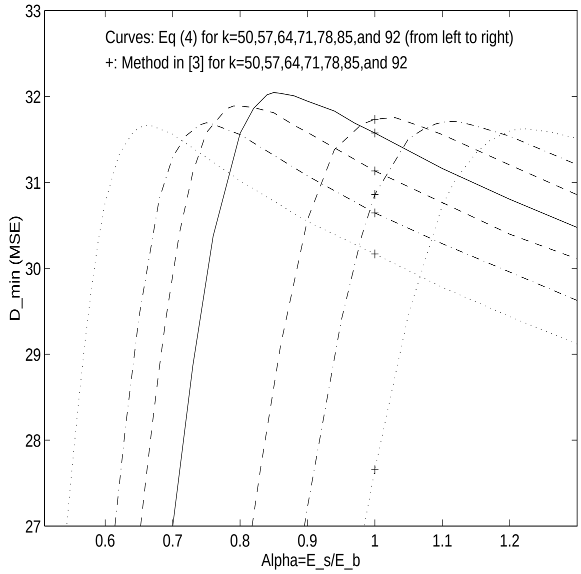

Figure 4: PSNR values for different BCH codes of length 127, for $p(E_b) = 10^{-2}$ and $R_0 = 0.25$ bpp.

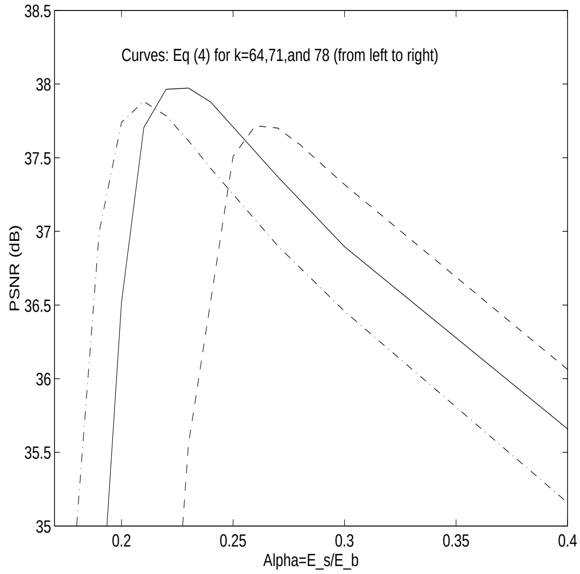

Figure 5: PSNR values for different BCH codes of length 127, for $p(E_b) = 10^{-6}$ and $R_0 = 0.25$ bpp.