# Progressive Image Communication over Binary Channels with Additive Bursty Noise *

Firouz Behnamfar†, Fady Alajaji†‡, and Tamás Linder†‡
† Department of Electrical and Computer Engineering
‡ Department of Mathematics and Statistics
Queen's University, Kingston, Ontario, Canada.

**Abstract**

A progressive method for transmission of images over a bursty noise channel is presented. It is based on discrete wavelet transform (DWT) coding and channel-optimized scalar quantization. The main advantage of the proposed system is that it exploits the channel memory and hence has superior performance over a similar scheme designed for the equivalent memoryless channel through the use of channel interleaving. In fact, the performance of the proposed system improves as the noise becomes more correlated, at a fixed bit error rate. Comparisons are made with other alternatives which employ independent source and channel coding over the fully interleaved channel at various bit rates and bit error rates. It is shown that the proposed method outperforms these substantially more complex systems for the whole range of considered bit rates and for a wide range of channel conditions.

## 1 Introduction

Justified by Shannon's separation principle [1], the traditional approach to data communication is to design "tandem systems" which are composed of independent source and channel coders. Being only asymptotically optimal, tandem systems often result in either too conservative or insufficient protection against channel errors. In multimedia communications, the decoder output would have a coarse resolution in the former case, while the latter case would result in poor quality.

It is known that in real world applications, with restrictions on delay and complexity, joint design of the source and channel coders results in superior performance. Several methods have been proposed for joint source-channel coding, which may be categorized as unequal error protection [2]-[7], channel-optimized scalar and vector quantization [8]-[12], optimization of index assignment [13, 14], and exploitation of the residual redundancy of the source coder via MAP decoding [15]-[19]. All of the above methods may be applied in image transmission, but the first two approaches appear to draw more attention in the literature.

Unequal error protection trades off source resolution and channel error protection to allocate the available bit rate. The amount of channel coding depends on the channel conditions, such as the bit error rate. Most of such strategies employ rate-compatible punctured convolutional (RCPC) codes for error protection [20], often with a measure of parity checking, such as cyclic redundancy check (CRC), for

*This work was supported in part by NSERC of Canada and PREA of Ontario.
Authors e-mail addresses: {`firouz, fady, linder`}`@mast.queensu.ca`

detection of errors in the channel-decoded sequence. In [2], entropy-coded subband coding is employed for compression. Packetization and RCPC channel coding are then used to transmit images over memoryless noisy channels. While the coefficients at each subband are given the same level of protection in [2], in [3] different rates of the RCPC code are used in each subband according to the sensitivity of the end-to-end distortion to each bit significance level. In [4], the strong image coder SPIHT introduced in [21] is applied for compression, together with packetization, check sum bits, RCPC channel coding and list Viterbi decoding [22]. An adaptive source and channel coding rate allocation scheme for finite state channels is examined in [5]. The work in [6] uses the Gilbert-Elliott channel model [23] to allocate the RCPC code rate to various blocks inside each image subband. The so-called "all-pass filtering" [24] is also applied to increase the peak signal to noise ratio (PSNR) by 3 dB. The Gilbert-Elliott model was also used in [7], but neither of the two latter coders exploit the channel memory.

Channel optimized vector quantization (COVQ) is another alternative for joint source-channel coding [8, 9]. In a typical implementation of this method, one may modify the nearest neighbor and centroid equations of the generalized Lloyd algorithm (GLA), taking the channel crossover probabilities into account, hence designing a codebook for each channel condition. After quantization, the binary indices are sent directly over the channel. The decoder simply uses the received noisy bits to find the transmitted indices. In [11], subband coding, all-pass filtering, and channel optimized scalar quantization (COSQ) are applied for image transmission over memoryless noisy channels. This method is similar in spirit to our approach, but our channel has memory and we use DWT-based subband coding which is substantially less complex due to reduced filter lengths. Our work extends and improves the results in [12], which uses DCT coding, specially at low bit rates.

One may encounter a combination of the above methods in the literature. For example, in [25], a COVQ is designed for a channel equivalent to the combination of an RCPC encoder, a memoryless channel, and an RCPC decoder.

An important property of many real life channels, including wireless channels,is their memory. Using an efficient interleaver tends to render such channels memoryless; however, even ideal interleavers result in delay and complexity. Furthermore, the resulting associated memoryless channel has a lower capacity than the original channel with memory (for the case of information stable channels [26]). Thus, it would be more beneficial to deal with channels with memory in a different way. In wireless image transmission where bandwidth is scarce, it is not practical to use two concatenated channel codes, such as convolutional or Turbo codes and Reed-Solomon codes, because the resolution of the received image would be too coarse, given a fixed bit rate. COVQ may be employed in such cases. The fact that explicit knowledge of channel block transition probabilities is a necessity for COVQ design restricts their use, because it is not always easy to derive closed form expressions for the transition probabilities. In [26], an alternative to the Gilbert-Elliott channel model is presented and the channel block probabilities are explicitly derived. This model is used in [27] to design COSQ and COVQ for generalized Gaussian as well as Gauss-Markov sources.

In this paper, we present an image compression method for the noisy channel with memory introduced in [26]. The algorithm is composed of two parts: a DWT-based transform coder and a channel optimized scalar quantizer which uses a *fixed* bit allocation table for all image tiles, as opposed to JPEG2000 and its coding engine EBCOT [28] where different bit allocations are provided for different blocks. One salient feature of this method is that its performance *improves* as channel noise becomes more correlated, making it attractive for wireless channels. Also, the rate can virtually be changed continuously and hence very efficient use of the available bit rate is possible. Moreover, it provides reasonable image quality at bit rates as low as 0.125 bpp and bit error rates as high as 0.1. Our system performs better than a similar system designed for a memoryless channel and used in series with an *ideal* interleaver, which increases delay. Also, it outperforms unequal error protection schemes which use scalar quantization, convolutional coding and ideal interleaving. The proposed method offers the following features.

- The channel interleaver is left out. This eliminates additional delay and, more importantly, exploits to a bigger extent the capacity of the channel with memory, which is larger than the capacity of the equivalent (interleaved) memoryless channel.

- Unlike unequal error protection schemes (e.g., [2]-[4], [6, 7]) the computational complexity of bit allocation is negligible.

- Unlike many other schemes, such as unequal error protection and [28], encoding is performed only once; no multi-stage source compression and subsequent channel coding is required.

- The proposed method is progressive. Beginning from the lowest resolution subband, the transmitter sends the data of the next resolution level every time.

Note that for now our objective is to demonstrate the advantages of COSQ over methods which use channel interleaving to suppress the effect of channel memory. We consider bursty channels with memory, for which complex schemes which use RCPC codes (such as [2] and [4]) might not be suitable.

## 2 COSQ-Based Image Coding for Bursty Noise Channels

*A. Structure*

Figure 1 shows the block diagram of the employed image coding system. Tiling is simply dividing the image into blocks of size $2^m \times 2^m$ where $m \geq 3$ is an integer.

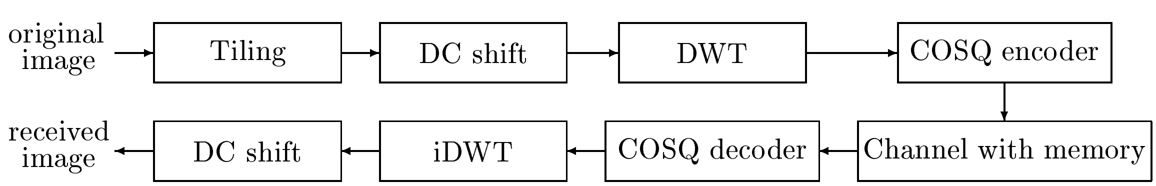


Figure 1: The structure of the proposed coding system.

It is known that larger tile sizes result in higher PSNR [29]; they however require more memory. The DC shifter takes a constant, say 128, off all entries in each tile. Subtracting the average value of each tile from its entries would result in higher PSNR; but it increases the amount of side information. Next, the DWT of each tile is extracted three times, every time on the lowest resolution subband of the previous resolution level. The schemes in [2]-[4], [6, 7, 11, 21] aim to exploit the intra-block dependencies by considering groups of coefficients which are expected to have high correlation and call them "sub-sources". Our method is different in that we use the inter-tile dependencies. If $c_{i,j}^{(k)}$ is the coefficient at row $i$ and column $j$ of the $k^{\text{th}}$ tile, our $(i,j)^{\text{th}}$ sub-source is $\{c_{i,j}^{(k)}\}_{k=1}^{T}$ where $T$ is the number of tiles. Depending on the available bit rate, some of the sub-sources are then normalized (to have a unit variance) and quantized, using a COSQ for channels with memory. The resulting bit-stream is sent directly over the channel. The receiver is simply the inverse of the transmitter.

For COSQ design, we need to know the distribution of the samples to be quantized. It is well-known that the distribution of the DWT coefficients of images approximately follows the generalized Gaussian distribution [11, 30], with a probability density function given by

$$f(x) = \frac{\alpha\eta(\alpha,\sigma)}{2\Gamma(1/\alpha)} \exp\{-[\eta(\alpha,\sigma)|x|]^{\alpha}\} \tag{1}$$

where $\eta(\alpha,\sigma) = \frac{1}{\sigma}\left(\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}\right)^{\frac{1}{2}}$ is the rate of decay, $\sigma^2$ is the variance, and $\Gamma(\cdot)$ is the Gamma function. For $\alpha$=1 and 2, the above yields the Laplacian and Gaussian distributions, respectively. For simplicity, we assume here that the sub-sources in all subbands have the Laplacian distribution and we quantize them using a COSQ trained for such a source with a unit variance.

*B. Model for Channel with Memory*

Based on [26], we use the Polya-contagion channel model which assumes that any noise sample depends only on the sum of the $M$ previous samples. The resulting noise process is a stationary ergodic Markov source of order $M$. If $X_i$, $Y_i$, and $Z_i$ represent the input, output, and noise in that order and $\oplus$ is addition modulo 2, the channel input-output relationship is described by $Y_i = X_i \oplus Z_i$. Assuming that the input and noise are independent, for $i \geq M$ we have, for any $e_{i-M}^{i-1} \in \{0,1\}^M$ (see [26]):

$$\Pr\{Z_i = 1 | Z_{i-M}^{i-1} = e_{i-M}^{i-1}\} = \Pr\left\{Z_i = 1 \,\middle|\, \sum_{j=i-M}^{i-1} Z_j = \sum_{j=i-M}^{i-1} e_j\right\} = \frac{\epsilon + \delta\sum_{j=i-M}^{i-1} e_j}{1 + M\delta}$$

where $\epsilon$ is the bit error rate (BER) and $\delta \geq 0$ controls the correlation coefficient of the noise given by $\frac{\delta}{\delta+1}$. The channel capacity (whose closed-form expression is derived in [26]) increases with $\delta$, showing that COVQ may achieve less distortion for channels with memory. This is supported by the simulation results in [27] for generalized Gaussian sources. If $\delta$ is set to zero, the noise process becomes memoryless and the channel reduces to a binary symmetric channel (BSC). Remark also that this model is less complex than the Gilbert-Elliott channel model and is completely specified with only three parameters.

### *C. COSQ Design*

If $d = d_H(\mathbf{x}, \mathbf{y})$ denotes the Hamming distance between the binary channel input block $\mathbf{x}$=$(x_1, \ldots, x_n)$ and the output block $\mathbf{y}$=$(y_1, \ldots, y_n)$, we have [26]:

- For $n \leq M$, $\mathrm{P}(\mathbf{y}|\mathbf{x}) = L(n, d, \epsilon, \delta)$, where

$$L(n, d, \epsilon, \delta) = \frac{\prod_{i=0}^{d-1}(\epsilon + i\delta)\prod_{i=0}^{n-d-1}(1 - \epsilon + i\delta)}{\prod_{i=0}^{d-1}(1 + i\delta)}.$$

- For $n > M$,

$$\mathrm{P}(\mathbf{y}|\mathbf{x}) = L(M, s_{M+1}, \epsilon, \delta)\prod_{i=M+1}^{n}\left[\frac{\epsilon + s_i\delta}{1 + M\delta}\right]^{e_i}\left[1 - \frac{\epsilon + s_i\delta}{1 + M\delta}\right]^{1-e_i}$$

  where $e_i = x_i \oplus y_i$ and $s_i = e_{i-M} + \cdots + e_{i-1}$.

The significance of the above formulas is that unlike many other channel models in the literature, they provide easy and computationally inexpensive tools to implement the GLA algorithm for noisy channels.

Various algorithms have been proposed for COVQ design among which we name the modified GLA initialized by simulated annealing [13], noisy channel relaxation [31], stochastic relaxation [32], and deterministic annealing [33]. After some experiments with the above algorithms, we decided to use the modified GLA with simulated annealing because of its computational efficiency. Unlike other methods, it computes the codebooks for rates as high as 9 bits per sample in a reasonable time.

### *D. Bit Allocation*

It is well known that in DWT subband coding, the end-to-end distortion is more sensitive to errors in the low resolution subbands. Therefore, when allocating bits to the sub-sources, the subbands at which they are located should be taken into account.

Usually, the distortion of sub-source $i$ is weighted by the $L_2$ norm of the wavelet basis functions of the subband to which it belongs, denoted by $w_i^2$. The end-to-end distortion is then written as $D = \sum_{i=1}^{S} w_i^2 d_i$, where $S$ is the number of sub-sources and $d_i$ is the distortion of sub-source $i$. Note that if two sub-sources are at the same subband, they are weighted identically. As often used in the literature, we use the mean squared error distortion measure and hence $d_i = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}(c_{m,n} - \hat{c}_{m,n})^2$ for a sub-source of size $M \times N$. The above weights are derived by approximation [28] and we use other weights which are based on heuristics and result in up to 0.2 dB overall gain in PSNR. Because of the way we built our sub-sources, the sub-sources may have different weights even within the same subband. Therefore, we represent the end-to-end distortion as

$$D = \sum_{i=1}^{L} \sum_{j=1}^{L} w_{i,j} d_{i,j}, \qquad d_{i,j} = \frac{1}{T} \sum_{k=1}^{T} (c_{i,j}^{(k)} - \hat{c}_{i,j}^{(k)})^2 \tag{2}$$

for a tile size of $L \times L$ and $T$ tiles. Our heuristic method is as follows: Considering a tile, we change one coefficient by a given number, compute the inverse DWT and evaluate the distortion. Having done this for all coefficients, we normalize the results by the largest absolute value to get the relative sensitivities. We repeated this experiment for various perturbations. While the results were nearly constant for perturbations larger than 10, we applied the average values in our study. The result for tile size 8×8 and the Daubechies irreversible 9/7 filters [29] is given in Table 1.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1.000000 | 0.320970 | 0.049902 | 0.054666 | 0.017026 | 0.013157 | 0.014377 | 0.011178 |
| 0.320970 | 0.096788 | 0.111598 | 0.123654 | 0.020076 | 0.015414 | 0.016787 | 0.013148 |
| 0.049902 | 0.111598 | 0.017467 | 0.021970 | 0.020997 | 0.016066 | 0.017496 | 0.013707 |
| 0.054666 | 0.123654 | 0.021970 | 0.022344 | 0.029752 | 0.023575 | 0.025070 | 0.019940 |
| 0.017026 | 0.020076 | 0.020997 | 0.029689 | 0.006893 | 0.005148 | 0.005430 | 0.004485 |
| 0.013157 | 0.015414 | 0.016066 | 0.023575 | 0.005148 | 0.003757 | 0.004615 | 0.003270 |
| 0.014377 | 0.016787 | 0.017496 | 0.025070 | 0.005430 | 0.004615 | 0.004894 | 0.003865 |
| 0.011178 | 0.013148 | 0.013707 | 0.019940 | 0.004485 | 0.003270 | 0.003865 | 0.002846 |

Table 1- Table of sensitivities for tile size $8 \times 8$.

We employed dynamic programming for bit allocation. In particular, we extended the work in [10] to the Markov channel and to the case where the overall distortion has different sensitivities to different sub-sources. The bit allocation problem is to minimize (2) subject to $\sum_{i=1}^{L} \sum_{j=1}^{L} r_{i,j} \leq B$ and $0 \leq r_{i,j} \leq r_{\max}$, where $r_{i,j}$ is the number of bits allocated to the $(i,j)^{\text{th}}$ sub-source and $B$ is the total number of bits available, so that the overall bit rate equals $B/L^2$ (neglecting the side information for now). $r_{\max}$ is the maximum number of bits which may be allocated to a sub-source. In this paper, we choose $r_{\max} = 9$ bits to have relatively small codebooks and fast encoding. Modeling the sub-sources as independent Laplacian sources, we can write each $d_{i,j}$ in (2) as $\sigma_{i,j}^2 d_L(r_{i,j})$ where $d_L(r_{i,j})$ is the distortion of a unit-variance Laplacian source quantized for a set of channel conditions ($i.e.$, $\epsilon, \delta, M$) and $\sigma_{i,j}^2$ is the variance of the $(i,j)^{\text{th}}$ sub-source. The problem now is to allocate the available bits to $L^2$ Laplacian sources, each with variance $w_{i,j}\sigma_{i,j}^2$, given the channel conditions. We use the algorithm in [10] to solve this problem which is guaranteed by [10] to achieve optimal bit allocation.

Note that $d_L(r_{i,j})$ is calculated off-line. Also, although $\sigma_{i,j}^2$ is image-dependent, it is not computed inside the algorithm. Indeed, the computational complexity of this algorithm is favorable compared to [3, 4, 6, 7, 28].

## 3 Simulation Results

We implemented the proposed image coder for the compression and transmission of gray-scale images over the contagion channel with $M = 1$ and tested it for the image

Lena (tests performed on other images such as Goldhill, Baboon, and Peppers gave results consistent with the Lena experiments). The simulation results of all tested systems are shown in Figure 2. We refer to our system as COSQ, followed by the $\delta$ of the channel it is designed for (*e.g.*, COSQ-5 and COSQ-10). COSQ-IL denotes the same system which uses an ideal channel interleaver, and hence it is designed for the memoryless BSC. Three other tandem systems are also considered. They employ scalar quantization, a very strong convolutional code and an ideal interleaver. The convolutional codes are selected from [22] and are all nonsystematic and have 64 states (6 memory elements per output bit). They are the strongest convolutional codes reported in [22] for this memory size, with $d_{\text{free}}$= 10, 14, and 20 for rates 1/2, 1/3, and 1/4, respectively. At the receiver, the maximum likelihood Viterbi algorithm is used. We refer to these systems as "CC 1/2 IL", "CC 1/3 IL", and "CC 1/4 IL" in that order. The best performance among the tandem systems is shown by "C best". This could represent the performance curve of a UEP system using an RCPC code with a very small mother code rate and a large puncturing period. The rate allocation algorithm for such a system is computationally prohibitive, but this curve is helpful for comparison purposes. As expected, when BER increases, "C best" begins with "C 1/2 IL", switches to "C 1/3 IL" at intermediate BER values and ends with "C 1/4 IL". As smaller tile sizes require less memory, they are attractive for hardware implementation; therefore, two tile sizes are considered, $8\times8$ and $64\times64$. Each test was repeated 50 times. The average PSNR is reported for various total bit rate and BER values. The PSNR is defined as $10\log_{10}\frac{255^2}{D}$, where $D$ is the mean square error between the original and decoded images.

We clearly observe that the performance curve of the COSQ-IL is always lower than the curves of COSQ-5 and COSQ-10. This shows that it is preferable to exploit the channel memory rather than destroying it. The COSQ-based method for correlated noisy channels is also better than the interleaver-based tandem systems when noise is correlated. Indeed, substantial coding gains are observed for the whole range of the channel BER. As mentioned earlier, this is because the COVQ system exploits the larger capacity of the channel with memory for all BERs and the fact that, at high BERs, the RCPC decoders of the tandem schemes fail to correct all channel errors, causing an error propagation effect.

Note that the results derived for the tandem schemes are indeed an upper-bound to their performance. In reality, no ideal interleaver exists. Tandem systems are substantially more complex and introduce considerable delay. Moreover, they are very sensitive to channel memory; a performance degradation as high as 10.5 dB was observed in "C best" when we applied the correlated noise with $\delta = 5$ and $M = 1$ at total bit rate of 1 bpp.

Throughout this work, we considered binary channels with memory which model physical channels used in conjunction with hard-decision demodulation. Future work might address the design of efficient COSQ-based image coding schemes for soft-decision demodulated channels with memory. It is expected that additional coding gains can be obtained via the use of the channel soft-decision information; this was indeed observed in [34, 35] for the case of ideal Gaussian sources.

## References


[1] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," in *IRE Nat. Conv. Rec.*, Pt. 4, pp. 142-163, 1959.

[2] N. Tanabe and N. Farvardin, "Subband image coding using entropy-coded quantization over noisy channels," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 926-943, June 1992.

[3] M. J. Ruf and J.W. Modestino, "Rate-distortion performance for joint source and channel coding of images," in *Proc. ICIP*, vol. II, pp. 77-80, 1995.

[4] P. G. Sherwood and K. Zeger, "Progressive image coding on noisy channels," in *Proc. DCC*, Snowbird, UT, Mar. 1997.

[5] H. Jafarkhani, P. Ligdas and N. Farvardin, "Adaptive rate allocation in a joint source/channel coding framework for wireless channels," in *Proc. IEEE Veh. Technol. Conf.,* vol. 1, pp. 492-296, 1996.

[6] J. Cai and C. W. Chen, "Robust joint source-channel coding for image transmission over wireless channels," *IEEE. Trans. Circuits Syst. Video Technol.*, vol. 10, pp. 962-966, Sept. 2000.

[7] J. Lu, A. Nosratinia and B. Aazhang, "Progressive source-channel coding of images over bursty error channels," in *Proc. ICIP*, Chicago, IL., Oct. 1998.

[8] H. Kumazawa, M. Kasahara, and T. Namekawa, "A construction of vector quantizers for noisy channels," *Electronics Eng. in Japan*, vol. 67-B, pp. 39-47, Jan. 1984.

[9] N. Farvardin and V. Vaishampayan, "On the performance and complexity of channel optimized vector quantizers," *IEEE Trans. Inform. Theory.*, vol. 37, pp. 155-160, Jan. 1991.

[10] V. Vaishampayan and N. Farvardin, "Optimal block cosine transform image coding for noisy channels," *IEEE Trans. Commun.*, vol. 38, pp. 327–336, Mar. 1990.

[11] Q. Chen and T. R. Fisher, "Image coding using robust quantization for noisy digital transmission," *IEEE Trans. Image Processing*, vol. 7, pp. 496-505, Apr. 1998.

[12] J. Cheng and F. Alajaji, "Channel optimized quantization of images over bursty channels," in *Proc. The 1997 Canadian Workshop on Inform. Theory*, June 1997.

[13] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Trans. Inform. Theory*, vol. 36, pp. 799-809, July 1990.

[14] K. Zeger and A. Gersho, "Pseudo-Gray coding," *IEEE Trans. Commun.*, vol. 38, pp. 2147-2158, , Dec. 1990.

[15] K. Sayood and J. C. Borkenhagen, "Use of residual redundancy in the design of joint source/channel coders," *IEEE Trans. Commun.*, vol. 39, pp. 838-846, June 1991.

[16] W. Xu, J. Hagenauer and J. Hollmann, "Joint source-channel coding using the residual redundancy in compressed images," in *Proc. ICC*, Dallas, TX., June 1996.

[17] F. Alajaji, N. Phamdo and T. Fuja, "Channel codes that exploit the residual redundancy in CELP-encoded speech," *IEEE Trans. Speech Audio Processing,* vol. 4, pp. 325-336, Sept. 1996.

[18] P. Burlina and F. Alajaji, "An error resilient scheme for image transmission over noisy channels with memory," *IEEE Trans. Image Processing*, vol. 7, pp. 593-600, Apr. 1998.

[19] B. Srinivas, R. Ladner, M. Azizoğlu and E. Riskin, "Progressive transmission of images using MAP detection over channels with memory," *IEEE Trans. Image Processing*, vol. 8, pp. 462-475, Apr. 1999.

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

[21] 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.

[22] R. Johannesson and K. Zigangirov, *Fundamentals of Convolutional Coding.* New York, NY: IEEE Press, 1999.

[23] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, pp. 1253-1265, Sept. 1960.

[24] C. J. Kuo and C. S. Huang, "Robust coding technique-transform encryption coding for noisy communications," *Optical Engineering* vol. 32, pp. 150-156, Jan. 1993.

[25] A. Goldsmith and M. Effros, "Joint design of fixed-rate source codes and multiresolution channel codes," *IEEE Trans. Commun.*, vol. 46, pp. 1301-1312, Oct. 1998.

[26] F. Alajaji and T. Fuja, "A communication channel modeled on contagion," *IEEE Trans. Inform. Theory*, vol. 40, pp. 2035-2041, Nov. 1994.

[27] N. Phamdo, F. Alajaji, and N. Farvardin, "Quantization of memoryless and Gauss-Markov sources over binary Markov channels," *IEEE Trans. Commun.*, vol. 45, pp. 668-675, June 1997.

[28] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Trans. Image Processing*, vol. 9, pp. 1158-1170, July 2000.

[29] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 image coding system: an overview," *IEEE Trans. Consumer Electron.*, vol. 46, pp. 1103-1127, Nov. 2000.

[30] S. M. LoPresto, K. Ramchandran and M. T. Orchard, "Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework," in *Proc. DCC*, Snowbird, UT, Mar. 1997.

[31] S. Gadkari and K. Rose, "Robust vector quantizer design by noisy channel relaxation," *IEEE Trans. Commun.*, vol. 47, pp. 1113-1116, Aug. 1999.

[32] K. Zeger, J. Vaisey, and A. Gersho, "Globally optimal vector quantizer design by stochastic relaxation," *IEEE Trans. Signal Processing*, vol. 40, pp. 310-322, Feb. 1992.

[33] D. Miller and K. Rose, "Combined source-channel vector quantization using deterministic annealing," *IEEE Trans. Inform. Theory*, vol. 42, pp. 347-356, Feb./Mar./Apr. 1994.

[34] F. Alajaji and N. Phamdo, "Soft-decision COVQ for Rayleigh-fading channels," *IEEE Commun. Let.*, vol. 2, pp. 162-164, June 1998.

[35] M. Skoglund and P. Hedelin, "Hadamard-based soft decoding for vector quantization over noisy channels," *IEEE Transactions on Information Theory*, vol. 45, pp. 515–532, Mar. 1999.

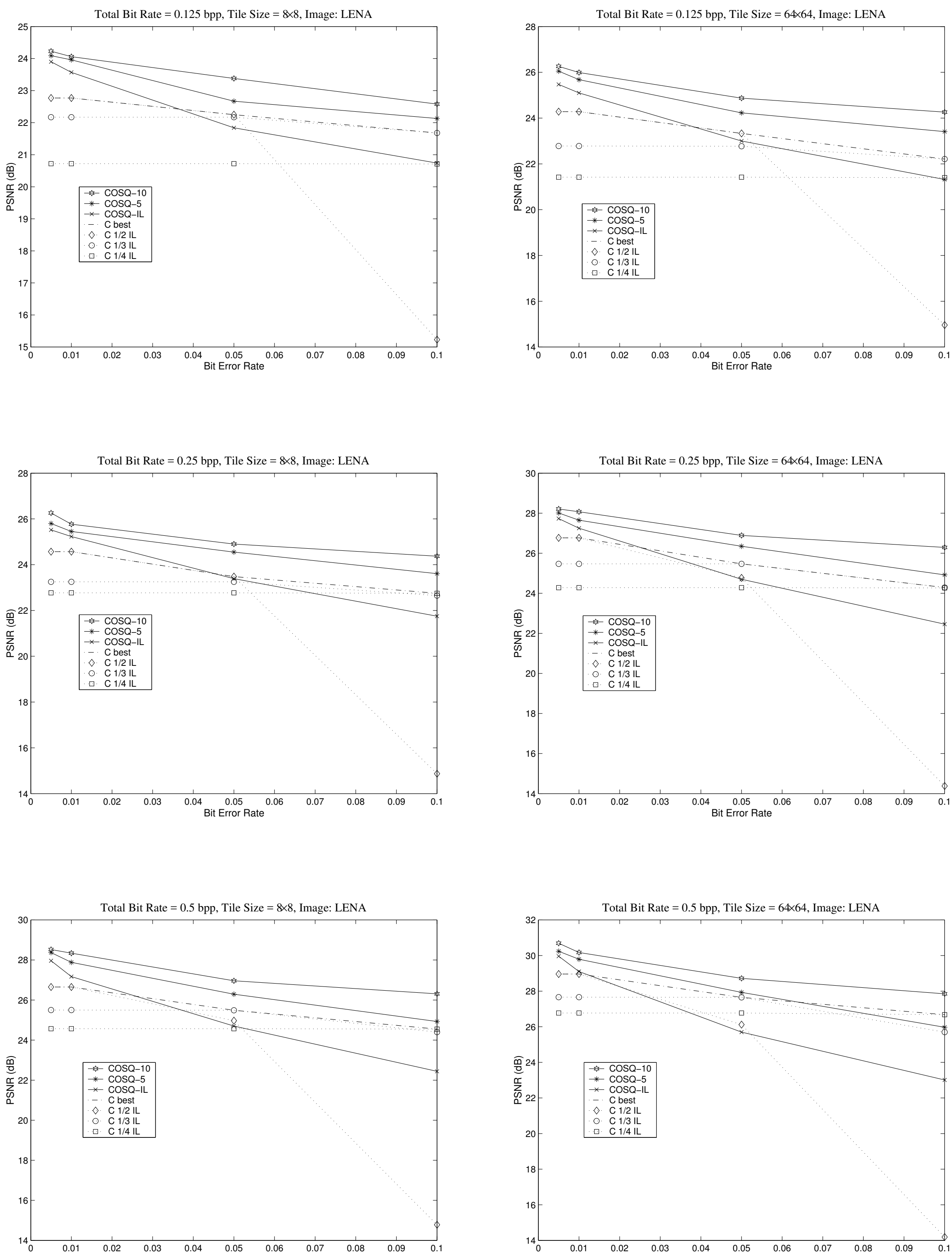


Figure 2: Performance of the implemented image coders for various channel conditions and bit rates.