Mathematics and Engineering Communications Laboratory

Technical Report

Channel Optimized Quantization of Images over Binary Channels with Memory

J. Cheng and *F. Alajaji*

April 1999

Channel Optimized Quantization of Images over Binary Channels with Memory*

Julian Cheng[†] and *Fady Alajaji*[‡†]

[†] Department of Electrical and Computer Engineering
Queen's University, Kingston, Ontario, Canada, K7L 3N6

[‡] Department of Mathematics and Statistics
Queen's University, Kingston, Ontario, Canada, K7L 3N6

fady@polya.mast.queensu.ca

Abstract

This paper addresses the technique of joint source-channel coding for the efficient and reliable transmission of compressed images without the use of channel error-control codes over noisy channels with memory. The channel used is a binary channel with additive Markov noise. The proposed DCT-based system consists of a channel-optimized quantization scheme that exploits the channel memory by incorporating the characteristics of the correlated noise in the quantizer design. Experimental results show that this simple system - which employs a fixed zonal coding bit allocation technique - provides significant objective and subjective improvements over traditional tandem systems designed for the fully interleaved channel, especially during bad channel conditions. Performance gains are also observed over recent MAP-detection based joint source-channel coding schemes. The loss of optimality due to the use of the fixed zonal coding bit allocation method is also examined. The loss is shown to be small for various images; this suggests that a reduction in complexity and bandwidth requirements can further be achieved.

Keywords: Joint source-channel coding, robust quantization, image coding, channels with memory.

---

* This work is supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada. Parts of this paper were presented at the *Canadian Workshop on Information Theory*, Toronto, Canada, June 1997.

I Introduction

Traditionally, source and channel coding has been guided by Shannon's separation principle [23], which states that the source and channel coding operations can be designed independently from each other (in tandem) without loss of asymptotic optimality. However, in practice, a tandem coding system is constrained by the encoder/decoder delay and complexity. This drawback has motivated many researchers to investigate a joint source-channel approach as an alternative to the tandem system. Recently, joint source-channel coding has received considerable interest, particularly with regards to the design of quantization systems for noisy channels (e.g., [2]-[12], [14]-[20], [22], [24]-[29], [31], [33]).

Combined source-channel coding is one form of joint source-channel coding. In a combined source-channel coding system, the characteristics of both the source and the channel are incorporated into the design of a single code. Thus, this single code plays a dual role; it performs data compression while being error resilient at the same time. One salient feature of combined source-channel coding is that the quantizer trades extra quantization distortion for smaller channel distortion such that the overall distortion is minimized.

The vast majority of the previous work on joint source-channel coding assumes a memoryless channel model. The assumption of this channel model is too simplistic for most practical wireless communication channels which often exhibit memory. Traditionally, interleaving techniques are used to render such channels memoryless [19]. However, the resulting associated memoryless channel is known to have a lower capacity (for the case of information stable channels [1]). Furthermore, interleaving/de-interleaving introduces additional delay into the system. More recently, an effort has been directed for the development of joint source-channel coding methodologies that exploit the statistical structure of channels with memory instead of "destroying it" via interleaving (e.g., [2], [5], [20], [26], [27], [33]).

In this work, we investigate the problem of the efficient compression and robust communication of still images over noisy channels with memory. The channel model considered is a binary channel with additive Markov noise, where the noise source is generated via the recently investigated Polya contagion urn scheme [1]. We first propose and implement a channel optimized scalar quantization (COSQ) system for the transmission of discrete cosine transform (DCT) coded images. The statistics of the Markov noise are incorporated in the design of our COSQ system, resulting in a system that exploits the (intra block) memory of the channel [20]. The COSQ system is designed assuming a Gaussian source distribution for the DC coefficients and a Laplacian source distribution for the AC coefficients [21]. Experimental results demonstrate that this system – which utilizes a fixed (global) bit allocation table for the DCT coefficients – significantly outperform traditional tandem systems designed for the interleaved (memoryless) channel. Performance comparisons with a recent maximum-a-posteriori (MAP) detection based method studied in [5] are also provided. We next investigate the optimal bit assignment problem for the DCT coefficients in the COSQ system and evaluate the performance loss resulting from using a fixed allocation table. We also implement an alternative channel optimized vector quantization (COVQ) scheme that is applied directly on the image (without DCT coding) before transmission over the Markov channel.

In previous related works, Burlina and Alajaji developed a sequence MAP-detection approach that exploits the residual redundancy in images transmitted over binary channels with additive Markov noise in [5]. A similar MAP-detection technique was studied by Srinivas *et al.* in [27] for the progressive transmission of images over the Gilbert-Elliott channel. In [20], a COVQ scheme for the transmission of (ideal) stationary memoryless generalized Gaussian and Gauss-Markov sources over binary Markov channels was investigated. This paper applies and extends the previous work in [20] to the problem of the efficient quantization and error resilient transmission of (inherently non-stationary) image sources over channels with memory. The Polya-contagion Markov channel model is conveniently employed since it offers an interesting and less complex alternative to the Gilbert model [1, 20], and so that adequate comparisons with the MAP-detection system in [5] can be made.

The rest of this paper is organized as follows. In Section II, the Markov channel model is introduced. A general COVQ scheme that is optimized for the Markov channel is briefly described in Section III. In Section IV, a simple DCT-based COSQ system is proposed along with simulation results for various images. Comparisons are made against two tandem coding systems and other recently proposed joint source-channel coding systems. In Section V, the performance loss due to the use of non-optimal bit allocation tables for the DCT coefficients is assessed. The effect of mismatch conditions in the channel parameters on the system performance is examined in Section VI. In Section VII, an alternative COVQ system that is directly applied on the image pixels is presented and evaluated. Finally, concluding remarks are stated in Section VIII.

II Channel Model

One large class of channels with memory is the class of finite-state channels which includes the Gilbert-Elliott model [19] and is often used to model many realistic channels such as fading channels and channels with inter-symbol interference (ISI). In this paper, we consider a more explicit family of channels with memory, where the channel memory is exhibited via an additive Markov noise process. More specifically, we consider a binary channel with additive noise described by

$$Y_i = X_i \oplus Z_i \tag{1}$$

for $i = 1, 2, 3, \ldots$, where $\oplus$ represents modulo 2 addition, and X_i, Z_i, and Y_i are the channel input, noise and output respectively. The input and noise sequences are assumed to be independent of each other. The noise process $\{Z_i\}_{i=1}^{\infty}$ is generated by the finite-memory Polya contagion urn model described in [1]. The resulting noise source $\{Z_i\}_{i=1}^{\infty}$ is a stationary ergodic Markov process of order M, i.e., for $i \geq M+1$,

$$Pr\{Z_i = e_i | Z_{i-1} = e_{i-1}, \ldots, Z_1 = e_1\} = Pr\{Z_i = e_i | Z_{i-1} = e_{i-1}, \ldots, Z_{i-M} = e_{i-M}\} \tag{2}$$

where $e_i \in \{0, 1\}$. Furthermore, $\{Z_i\}$ depends only on the sum of the M previous noise samples, and the noise transition probability is given by, for $i \geq M+1$,

$$Pr\{Z_i = 1 | Z_{i-M} = e_{i-M}, \ldots, Z_{i-1} = e_{i-1}\} = \frac{\epsilon + (\sum_{j=i-M}^{i-1} e_j)\delta}{1 + M\delta}, \tag{3}$$

where $e_j \in \{0,1\}$, for $j = i-M, \ldots, i-1$, and ϵ is the channel bit error rate (BER), which determines the marginal distribution of the noise process,

$$Pr\{Z_i = 1\} = \epsilon = 1 - Pr\{Z_i = 0\}. \tag{4}$$

The non-negative parameter δ determines the amount of correlation in $\{Z_i\}$ and it is a measure of the burstiness within the noise samples. The higher the noise correlation δ is, the more bursty the channel becomes. The correlation coefficient of the noise process is given by $\frac{\delta}{1+\delta}$ [1]. Note that when $\delta = 0$, the channel model reduces to the (memoryless) binary symmetric channel (BSC) with cross-over probability ϵ. We further observe that the above channel can be entirely described with only three parameters (ϵ, δ, and M). This channel model offers a possible alternative to finite-state channels such as the Gilbert-Elliott noise model.

In this paper, we only consider the first-order ($M = 1$) Markov noise process case; the noise transition probability can be found as

$$Pr\{Z_i = 1 | Z_{i-1} = e\} = \frac{\epsilon + e\delta}{1+\delta}, \tag{5}$$

where $e \in \{0,1\}$, or written in matrix form, the state transition probability becomes

$$\begin{bmatrix} Q(0|0) & Q(1|0) \\ Q(0|1) & Q(1|1) \end{bmatrix} = \begin{bmatrix} \frac{1-\epsilon+\delta}{1+\delta} & \frac{\epsilon}{1+\delta} \\ \frac{1-\epsilon}{1+\delta} & \frac{\epsilon+\delta}{1+\delta} \end{bmatrix}. \tag{6}$$

The capacity of this first-order Markov channel can be calculated as

$$\begin{aligned} C &= \lim_{n\to\infty} \max_{p(x^n)} \frac{1}{n} I(X^n; Y^n) & (7) \\ &= 1 - H(Z_2|Z_1) & (8) \\ &= 1 - \left[(1-\epsilon) h_b\left(\frac{\epsilon}{1+\delta}\right) + \epsilon h_b\left(\frac{\epsilon+\delta}{1+\delta}\right) \right], & (9) \end{aligned}$$

where $X^n \triangleq (X_1, X_2, \ldots, X_n)$, $Y^n \triangleq (Y_1, Y_2, \ldots, Y_n)$ and $H(Z_2|Z_1)$ is the entropy rate of the first-order Markov noise process, and $h_b(\cdot)$ is the binary entropy function. Note that the capacity of this channel model is monotonically increasing with δ (for fixed ϵ) and is monotonically decreasing with ϵ (for fixed δ). When the noise correlation δ increases, the channel becomes more bursty; in the extreme case when $\delta \to \infty$, the noise process becomes deterministic ($H(Z_2|Z_1)$ goes to zero) and the channel becomes noiseless. When the noise correlation $\delta = 0$, we obtain (as expected) the capacity expression of the BSC

$$C = 1 - h_b(\epsilon). \tag{10}$$

For an input block $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ and an output block $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$, the block channel transition probability matrix $Pr\{\mathbf{Y} = \mathbf{y}|\mathbf{X} = \mathbf{x}\}$ is given by [1], for $n \geq 2$,

$$\begin{aligned} Pr\{\mathbf{Y} = \mathbf{y}|\mathbf{X} = \mathbf{x}\} &\stackrel{\triangle}{=} Q(\mathbf{y}|\mathbf{x}) && (11) \\ &= Pr\{\mathbf{Z} = \mathbf{e}\} && (12) \\ &= Pr(Z_1 = e_1) \prod_{i=2}^{n} \left[\frac{\epsilon + e_{i-1}\delta}{1+\delta}\right]^{e_i} \left[\frac{(1-\epsilon) + (1-e_{i-1})\delta}{1+\delta}\right]^{1-e_i} && (13) \end{aligned}$$

where $e_i = x_i \oplus y_i, i = 1, 2, \ldots, n$. The above channel block distribution will be used later in designing channel optimized vector quantizers (COVQs).

III Channel Optimized Quantizer Design

In this section, we briefly formulate the problem of designing channel optimized vector quantizers (COVQ) for a channel with memory. The ensuing description of the COVQ design follows [20] [10]. Assume that the source to be encoded is a real-valued, stationary, and ergodic process $\{X_t; t = 0, 1, \ldots\}$ with zero mean and unit variance. The source is encoded with a k-dimensional, N-output level vector quantizer and the output of the VQ is transmitted over the binary Markov channel with input and output alphabets $\mathcal{J} = \{1, 2, \ldots, N\}$. The number of output levels N is usually taken values of power of 2 so there are $n = \log_2 N$ bits to describe each sample or level. The overall encoder-decoder operation can be decomposed into three separate mappings, namely the encoding mapping, $\gamma : \mathbf{R}^k \to \mathcal{J}$, the channel index mapping, $b : \mathcal{J} \to \mathcal{J}$, and the decoding mapping, $g : \mathcal{J} \to \mathbf{R}^k$.

The encoding mapping, γ is described in terms of the partition $\mathcal{P} = \{S_1, S_2, \ldots, S_N\}$ of the k-dimensional Euclidean space $\mathbf{R}^k$ according to $\gamma(\mathbf{x}) = i, \quad \text{if} \quad \mathbf{x} \in S_i, \ i \in \mathcal{J},$

where $\mathbf{x} = (x_1, x_2, \ldots, x_k)$ is a source output vector consisting of k source samples. The channel index mapping, b is a one-to-one mapping, which assigns the encoder output i an index $i' = b(i) \in \mathcal{J}$ and index i' is sent over the binary Markov channel. The channel is characterized by the channel transition probability $p(j|i')$ denoting the probability that the index j is received given that i' is transmitted and this transition probability is provided in equation (13). Finally, the decoding mapping is described in terms of the codebook $\mathcal{C} = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N\}$ according to $g(j) = \mathbf{y}_j, \quad j \in \mathcal{J}$.

We denote the distortion caused by representing the source vector $\mathbf{x}$ by a codeword $\mathbf{y}$ as $d(\mathbf{x}, \mathbf{y})$. The average overall distortion per source sample $D(\mathcal{P}, \mathcal{C})$ is described by

$$D(\mathcal{P}, \mathcal{C}) = \frac{1}{k} \sum_{i=1}^{N} \int_{S_i} p(\mathbf{x}) \left\{ \sum_{j=1}^{N} p(j|i) d(\mathbf{x}, \mathbf{y}_j) \right\} d\mathbf{x}. \tag{14}$$

where $p(\mathbf{x})$ is the k-fold probability density function of the source. Here we have dropped the index mapping b since it is insignificant in the COVQ design [10]. The encoding rate is given by

$$R = \frac{1}{k} \log_2 N \quad \text{bits/sample.} \tag{15}$$

Using the square-error distortion criterion, it can be shown [15] [10], that given the codebook $\mathcal{C}$ the optimal partition set can be expressed as

$$S_i^* = \left\{ \mathbf{x} : \sum_{j=1}^{N} p(j|i) ||\mathbf{x} - \mathbf{y}_j||^2 \leq \sum_{j=1}^{N} p(j|l) ||\mathbf{x} - \mathbf{y}_j||^2, \quad \forall l \right\} \quad i \in \mathcal{J}. \tag{16}$$

Similarly, the optimal codebook given the partition set is

$$\mathbf{y}_j^* = \frac{\sum_{i=1}^{N} p(j|i) \int_{S_i} \mathbf{x} p(\mathbf{x}) d\mathbf{x}}{\sum_{i=1}^{N} p(j|i) \int_{S_i} p(\mathbf{x}) d\mathbf{x}} \quad j \in \mathcal{J}. \tag{17}$$

The COVQ design procedure is a straightforward extension of the LBG-VQ design algorithm. The algorithm starts with an initial codebook, $\mathcal{C}^{(0)}$ to find the optimal partition set $\mathcal{P}^{(1)}$ using equation (16). With this newly computed $\mathcal{P}^{(1)}$, it uses equation (17) to update the optimal codebook $\mathcal{C}^{(1)}$. This process continues until the relative decrease in the average distortion is less than a specified threshold and the algorithm converges to a locally optimal solution. To obtain a channel optimized scalar quantizer (COSQ), we can simply take $k = 1$.

IV A DCT-COSQ Image Transmission System

A. System Description

In this subsection, we propose a DCT-based combined source-channel coding system to transmit grey-level images over the binary channel with additive Markov noise described in Section II. The block diagram of our proposed system is illustrated in Figure 1.

A grey-level (8 bpp or 256 levels) image is first subdivided into 8×8 blocks and transformed via the forward discrete cosine transform (FDCT) similar to the JPEG standard [32]. After proper normalization, higher frequency DCT coefficients are zonally masked out since they are relatively insensitive to the human visual system. Ideally, one desires to arrange the inter-block DC (or even AC) coefficients in a zigzag sequence similar to the JPEG standard [32]. It is well known that the inter-block DC coefficients are highly correlated (correlation coefficient, $\rho = 0.977$ for Lena and $\rho = 0.993$ for Baboon [13]). By arranging them in a zigzag fashion, it becomes desirable to exploit the memory within these transform coefficients by using channel optimized vector quantizers (COVQ). Unfortunately such an approach is not feasible for a coding system with high quantization rates, since the computational complexity and memory requirements grow exponentially with k and R. For example, by taking $k = 2$ and $R = 8$ for the DC coefficient, the size of the codebook is 2^{16}. We will therefore resort to the use of a COSQ system for the compression and transmission of the DCT coefficients.

Both DC and AC coefficients are quantized via a bank of channel optimized scalar quantizers (COSQ). *Fixed* bit allocation tables are used for each 8×8 image coefficient block. Since the DC coefficient (the coefficient with zero frequency) contains most of the energy in each image block, it is quantized with an 8-bit rate quantizer; as for the AC coefficients, they are quantized at rates that correspond to their level of activities. After quantization, the indices of the quantization level are coded via a natural binary code (NBC) and sent over the binary Markov channel. At the receiver end, they are decoded (instantaneously) and the reconstructed image is obtained through the inverse discrete cosine transform (IDCT). A bank of COSQs are designed off line using the method described in Section III. Here, we have assumed *a priori* knowledge of the channel conditions and the statistics of the quantizer input. The source distributions are assumed to be Gaussian for DC coefficients and Laplacian for all the AC coefficients [21]. We next evaluate the performance of the proposed system under various channel conditions.

B. System Performance and Comparisons

Experimental results for our proposed system indicate that large improvement over usual tandem schemes, which employ interleaving/de-interleaving and are designed for the noiseless channel, can be achieved. We performed the experiments on several images. To avoid exhaustive listings, we only present numerical results for Lena (512×512). In Tables 1-3, the average peak sgnal-to-noise ratio (PSNR) values of the reconstructed Lena are displayed for various values of the channel correlation δ, BER ϵ, and overall operational rate in bits per pixel (bpp). The objective measure PSNR (in dB) is defined as

$$PSNR = 10 \log_{10} \frac{255^2}{\mathbf{E}\{(X_{ij} - \hat{X}_{ij})^2\}} \tag{18}$$

where X_{ij} and $\hat{X}_{ij}$ are, respectively, the transmitted and reconstructed ij^{th} image pixel.

All simulation results were obtained by averaging over 25 experiments, and it was observed that the PSNR values vary very little from experiment to experiment. Three overall operational rates were used with fixed bit allocation tables listed in Table 4. The total number of bits B used for each 8×8 image block are 76, 58, and 24 bits (respectively yielding rates equal to 1.19, 0.9, and 0.375 bpp). The bit allocation table for the 1.19 bpp system is adopted directly from [30]; and in the 0.375 bpp system, only the first three transform coefficients are preserved and given the highest level of protection. By fixed bit allocation table, we mean that we apply the same table *globally* for *any* image under *any* channel condition. The advantage of using a fixed bit allocation method over using an adaptive optimal bit allocation technique is that the former one does not require overhead information. This results in a reduction of the encoder/decoder complexity and the bandwidth requirement of the overall system. The results obtained via this simple approach are not optimal. In the next section, we will study the loss of optimality of using such fixed bit allocation tables.

The performance results for the decoded image Lena are presented in Tables 1-3. In these tables, COSQ denotes our proposed scheme. Two reference tandem systems are next implemented for comparison purposes. The first reference system, denoted by SQ-IL, employs a Lloyd-Max quantizer for the DCT coefficients, followed with an NBC codeword assignment over the interleaved Markov channel. In this case, we assume that the Markov channel has been rendered memoryless (i.e. $\delta = 0$) via an ideal interleaver. Note that this system has the same quantization rate as our proposed COSQ system. The second reference system, denoted by CC-IL, employs a Lloyd-Max quantizer followed by channel coding on the codebook indices sent over the ideally interleaved (and thus memoryless) Markov channel. In order to achieve a fair comparison, we maintain an identical overall operating bit rate (bandwidth) as in our proposed scheme. Consequently, this system has a lower quantization rate. In essence, the CC-IL system trades less channel error for greater quantization error with respect to the SQ-IL system. The DCT coefficient are quantized according to bit allocation table provided in Table 5. The channel code used here is a rate-1/2 convolutional code with constraint length $K = 4$, $d_{free} = 6$ and generator polynomials $\mathbf{g}_0 = [1101]$ and $\mathbf{g}_1 = [1111]$ [34]. At the receiver, maximum-likelihood Viterbi decoding is used. Note the channel code rate is chosen somewhat heuristically as we do not address the problem of optimal rate allocation between the source and channel codes in this work (cf the recent works in [12, 14, 28]).

As shown from the PSNR tables, the data in the $\epsilon = 0$ column represent the PSNR values for the compressed image Lena, in which the distortions are exclusively due to the quantization errors. The data in the $\delta = 0$ row represent the PSNR values for the BSC. It is apparent that the COSQ system outperforms both reference systems in all cases, especially in very noisy channel environments with high noise correlation. More specifically, the improvement over the SQ-IL tandem system is as high as 12 dB for $\epsilon = 0.1$ and $\delta = 10.0$ and rate equal to 1.19 bpp. It can be observed that both SQ-IL and CC-IL schemes are extremely sensitive to the channel BER (ϵ); their performance degrade very quickly as the channel gets noisier. In contrast, when ϵ increases, the performance of the COSQ system degrades slowly and this is particularly true for high channel correlation parameters.

We next compare our simulation results with other recently proposed joint source-channel coding systems for the same Markov channel. Tables 1 and 2 include the numerical results for MAP-UNC, MAP-UEP I and MAP-UEP II described in [5]. Briefly, MAP-UNC, which does not employ channel coding, utilizes a MAP detector that exploits the residual source redundancy as well as the channel memory. MAP-UEP I (or II) employs additional channel coding for the DC coefficients via an unequal error protection scheme. Comparing the entries in the tables, it is clear that our combined source-channel coding scheme consistently outperforms the above joint source-channel schemes especially for channels with high error rates and high noise correlations. One reason why these MAP-detection joint source-channel coding systems underperform with respect to the COSQ system is that because they uniformly apply the JPEG standard quantization matrices on all 64 DCT coefficients, which can be inherently poor for noisy channels. Their performance can be improved by replacing the JPEG matrices by appropriately chosen Loyd-Max quantizers*.

In the context of image coding, no final judgment can be made without a subjective performance measure. In Figure 2, we show the image Lena sent using our proposed COSQ system and the SQ-IL and CC-IL tandem systems under severe channel conditions ($\epsilon = 0.1$). It can be clearly seen that the COSQ system provides the best performance. Additional results for the images Goldhill compressed at a medium bit rate (0.9 bpp), and for Peppers compressed at a low bit rate (0.375 bpp) are displayed in Figures 3 and 4, respectively. We conclude that the COSQ system outperforms the reference tandem coding schemes both objectively and subjectively.

We also observe that for a fixed ϵ, the performance of the COSQ scheme designed for the channel with memory outperforms the COSQ scheme designed for the memoryless channel. This is illustrated in Figure 5 for the image Lena compressed at a rate of 1.19 bpp. For example, at $\epsilon = 0.1$, there is (on average) an extra 3.45 dB gain for a system with $\delta = 5.0$

*However, as observed in [20] for the case of ideal stationary sources, it is expected that our COSQ system will still outperform the MAP system (for moderate to high values of kR) since it is designed to minimize the mean squared error while the MAP scheme minimizes the sequence error probability.

over the interleaved system ($\delta = 0$). Finally, it can be remarked from the figure that the gain due to the channel memory becomes more significant as the channel becomes noisier (i.e., as the BER increases).

V Bit Allocation Mismatch

Bit allocation addresses the proper distribution of the available bits to the transform coefficients. It determines which coefficients should be kept for coding and transmission and how coarsely the retained coefficients should be quantized. Bit allocation is usually performed with either adaptive threshold coding or zonal coding. Threshold coding is an adaptive method which is specified in the JPEG image standard. This coding scheme is based on the fact that different image blocks have different spectral and statistical behaviors so that adaptive bit allocation methods should be used for each image sub-block. Often in threshold coding a quantization matrix is applied to the sub-block coefficient matrix and only the coefficients with magnitudes above a certain threshold are retained. Thus the bit allocation table will differ from sub-block to sub-block. In zonal coding, the locations for the coefficients which carry the most information are retained via a zonal mask which is applied globally for each image sub-block. However, the zonal bit allocation table is calculated adaptively for each input image. In the presence of channel noise, channel optimized zonal coding translates into adaptively computing the optimal bit allocation tables by considering the channel parameters. In contrast, a simple fixed bit allocation scheme applies the same bit allocation table for any image and under any channel condition. It results in reducing the encoder/decoder complexity and does not require transmitting any overhead information. Obviously, a system which employs a fixed bit allocation scheme will not be optimal. We herein evaluate the loss of optimality due to the incorporation of the fixed bit allocation method in our proposed system. This is achieved by comparing it to a similar system which employs an optimal bit allocation scheme that minimizes the overall distortion. Here, we have assumed that the exact optimal bit allocation tables are available at the decoder. The task of calculating the optimal bit allocation table in the context of the Markov channel can be carried out by following the integer programming method provided in [31]. A detailed description of the algorithm is presented in [7]. The typical optimal bit allocation matrices are presented in Table 6 for Lena at 1.19 bpp under various channel conditions. Several interesting observations can be made by close examination of the bit allocation matrices. In all cases, most of the bits are concentrated, as expected, on the low frequency coefficients. For fixed δ, when the channel gets noisier, the first few low frequency coefficients receive the highest protection with the maximum number of allowable bits (8 bits). For fixed ϵ, when the channel gets more bursty (high δ), the bit distribution slightly spreads out to the higher frequency DCT coefficients.

Simulation results are presented in Tables 7-9, where COSQ-OPT denotes the schemes using the optimal bit allocation table and COSQ-FIX denotes the fixed bit allocation schemes. The results clearly show that the loss of optimality by using the fixed bit allocation table is very minor. In most cases, the loss is only about 1.0 dB in PSNR; this gap narrows as the channel conditions deteriorate. A similar behavior is observed for various other images [7].

VI Mismatch in Channel Parameters

In the design of channel optimized quantizer, it is assumed that the knowledge of the source distribution and the channel conditions are known *a priori* for the quantizer design. In this work, we assumed that the DC and AC transform coefficients follow the Gaussian and Laplacian distributions respectively. The source mismatch is minor in most cases [21]. Channel mismatch results have been reported in [9] for the BSC and in [20] for the compression of ideal sources over binary Markov channel. The performance for the image Lena under various channel mismatch conditions are listed in Tables 10 and 11. Here, we have followed the notation used in [20], where ϵ_a and ϵ_d denote the actual and designed BER; δ_a and δ_d denote the actual and designed noise correlation parameters. As shown from the PSNR data, for fixed δ, our proposed COSQ system are relatively insensitive to the channel BER mismatch provided that $\epsilon_d \geq \epsilon_a$. For fixed ϵ, we observe that the system performance are not very sensitive to the mismatch of the noise correlation. We conclude that in general, it is better to overestimate the true parameters rather than to underestimate them.

VII A COVQ Image Transmission System

As a completion to our simulation studies, we present some experimental results by using a COVQ system that is directly applied on the pixels of the image Lena (without DCT coding) before transmission over the Markov channel. The dimension was chosen to be 4×2 pixels (or $k = 8$) and a rate of 1.0 bpp was maintained. A set of training images (Goldhill, Airplane, Tiffany, Peppers, and Sailboat) was used to obtain the empirical k-fold source distribution. Again the quantizers were assumed to be matched to both the source distribution and the channel conditions. All results were obtained over 25 experiments and the PSNR performance for Lena are shown in Table 12. In this table, VQ-IL denotes a tandem system with VQ (designed for noiseless channel) and transmission of the codebook indices over the interleaved Markov channel. Similar observations can be made as in the case of the DCT-based COSQ system. The COVQ system outperforms the VQ-IL scheme in all cases, especially at bad channel conditions. By comparing the COVQ PSNR values with those from the COSQ system with rate 0.9 bpp in Table 2, we remark that the performance of the COVQ system is quite similar. However, it should be emphasized that the COVQ system depends on the training image sequences and has a high encoding complexity, while the DCT-based COSQ system, which uses a fixed source distribution for the DC and AC coefficients, is image independent.

VIII Conclusions

In this paper, we propose a DCT-based combined source-channel coding system for the reliable communication of grey-level images over binary bursty channels. The system consists of a channel optimized quantization scheme that exploits the channel memory. Experimental results demonstrate considerable objective and subjective performance improvements over traditional tandem coding schemes and some other recently proposed joint source-channel coding systems. We also illustrate the benefits of using the knowledge about the channel memory by incorporating it into the quantizer design, as opposed to employing interleaving and designing a system for the memoryless channel. The loss of optimality due to the use of fixed zonal coding is also studied; it is observed that such loss is relatively small. This suggests the possibilities for reducing the system complexity/bandwidth. Finally, an alternative COVQ system directly applied on the image pixels (without DCT coding) was implemented. Future work may include the investigation of robust image quantizers that exploit the channel soft decision information [3].

References

[1] F. Alajaji and T. Fuja, "A Communication Channel Model on Contagion," *IEEE Trans. Inform. Theory*, vol. 40, pp. 2035–2041, Nov. 1994.

[2] F. Alajaji, N. Phamdo, N. Farvardin, and T. Fuja, "Detection of Binary Markov Sources over Channels with Additive Markov Noise," *IEEE Trans. Inform. Theory*, vol. 42, pp. 230–239, January 1996.

[3] F. Alajaji and N. Phamdo, "Soft-Decision COVQ for Rayleigh Fading Channels," *IEEE Communications Letters*, vol. 2, pp. 162–164, June 1998.

[4] E. Ayanoglu and R. M. Gray, "The Design of Joint Source and Channel Trellis Waveform Coders," *IEEE Trans. Inform. Theory*, Vol. 33, pp. 855-865, November 1987.

[5] 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, April 1998.

[6] Q. Chen and T. R. Fischer, "Image Coding Using Robust Quantization for Noisy Digital Transmission," *IEEE Trans. Image Processing.*, vol. 7, pp. 496–505, April 1998.

[7] J. Cheng, *Channel Optimized Quantization of Images over Binary Channels with Memory*, M.Sc. Thesis, Queen's University, Kingston, Ontario, 1997. Available at: *http://markov.mast.queensu.ca/.*

[8] N. Farvardin and V. Vaishampayan, "Optimal Quantizer Design for Noisy Channels: An Approach to Combined Source-Channel Coding," *IEEE Trans. Inform. Theory*, vol. 33, pp. 827-838, Nov. 1987.

[9] N. Farvardin, "A Study of Vector Quantization for Noisy Channels," *IEEE Trans. Inform. Theory*, vol. 36, pp. 799–809, July 1990.

[10] 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.

[11] A. Fuldseth and T. A. Ramstad, "Channel-Optimized Subband Video Coding for Channels with a Power Constraint," *Proc. IEEE Int. Conf. Image Processing*, Chicago, IL, 1997.

[12] A. Goldsmith and M. Effros, "Joint Design of Fixed-Rate Source Codes and Multiresolution Channel Codes," *IEEE Trans. Commun.*, Vol. 46, pp. 1301-1312, October 1998.

[13] R. M. Gray, *Digital Image Processing Course Notes*, Stanford University, 1996.

[14] H. Jafarkhani, P. Ligdas and N. Farvardin, "Adaptive Rate Allocation in a Joint Source/Channel Coding Framework for Wireless Channels," *Proc. IEEE Vehicular Technology Conf.,* vol. 1, Atlanta, GA, April 1996.

[15] H. Kumazawa, M. Kasahara and T. Namekawa, "A Construction of Vector Quantizers for Noisy Channels," *Electronics and Engineering in Japan*, vol. 67-B, pp. 39–47, Jan. 1984.

[16] A. Kurtenbach and P. Wintz, "Quantizing for Noisy Channels," *IEEE Trans. Commun. Technology*, vol. 17, pp. 291–302, Apr. 1969.

[17] D. J. Miller and M. Park, "A Sequence-Based Approximate MMSE Decoder for Source Coding over Noisy Channels Using Discrete Hidden Markov Models," *IEEE Trans. Commun.*, Vol. 46, pp. 222–231, February 1998.

[18] J. Modestino, D. Daut, and A. Vickers, "Combined Source-Channel Coding of Images Using the Block Cosine Transform," *IEEE Trans. Commun.*, vol. 29, pp. 1261–1274, Sept. 1981.

[19] M. Mushkin and I. Bar-David, "Capacity and Coding for the Gibert-Elliott Channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 1277–1290, Nov. 1989.

[20] 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.

[21] R. Reininger and J. D. Gibson, "Distribution of the Two-Dimensional DCT Coefficients for Images," *IEEE Trans. Commun.*, vol. 31, pp. 835–839, June 1983.

[22] K. Sayood, F. Liu and J. D. Gibson, "A Constrained Joint Source/Channel Coder Design," *IEEE J. Selected Areas Commun.,* Vol. 12, pp. 1584-1593, December 1994.

[23] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623-656, 1948.

[24] P. G. Sherwood and K. Zeger "Error Protection for Progressive Image Transmission Over Memoryless and Fading Channels," *IEEE Trans. Commun.*, vol. 46, pp. 1555–1559, Dec. 1998.

[25] M. Skoglund, "A Soft Decoder Vector Quantizer for a Rayleigh Fading Channel – Application to Image Transmission," *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing,* vol. 5, Detroit, May 1995.

[26] M. Skoglund, "Soft-Decoding for Vector Quantization over Noisy Channels with Memory," *IEEE Trans. Inform. Theory*, to appear, May 1999.

[27] B. S. Srinivas, R. E. Ladner, M. Azizoğlu, and E. A. Riskin, "Progressive Transmission of Images Using MAP Detection over channel with Memory," *IEEE Trans. Image Processing*, vol. 8, pp. 462–475, April 1999.

[28] M. Srinivasan and R. Chellappa, "Adaptive Source-Channel Subband Video Coding for Wireless Channels," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1830–1839, Dec. 1998.

[29] N. Tanabe and N. Farvardin, "Subband Image Coding Using Entropy-Coded Quantization over Noisy Channels," *IEEE J. Sel. Areas Commun.*, vol. 10, pp. 926–943, June 1992.

[30] A. M. Tekalp, *Digital Video Processing.* Upper Saddle River, NJ: Prentice-Hall, 1995.

[31] V. A. Vaishampayan and N. Farvardin, "Optimal Block Cosine Transform Image Coding for Noisy Channels," *IEEE Trans. Commun.*, vol. 38, pp. 327–336, Mar. 1990.

[32] G. K. Wallace, "The JPEG Picture Compression Standard," *Commun. ACM*, pp. 31–43, Apr. 1991.

[33] H. S. Wang *Finite-State Modeling, Capacity, and Joint Source/Channel Coding for Time-Varying Channels*, PhD dissertation, Rutgers University, NJ, 1992.

[34] S. B. Wicker *Error Control Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

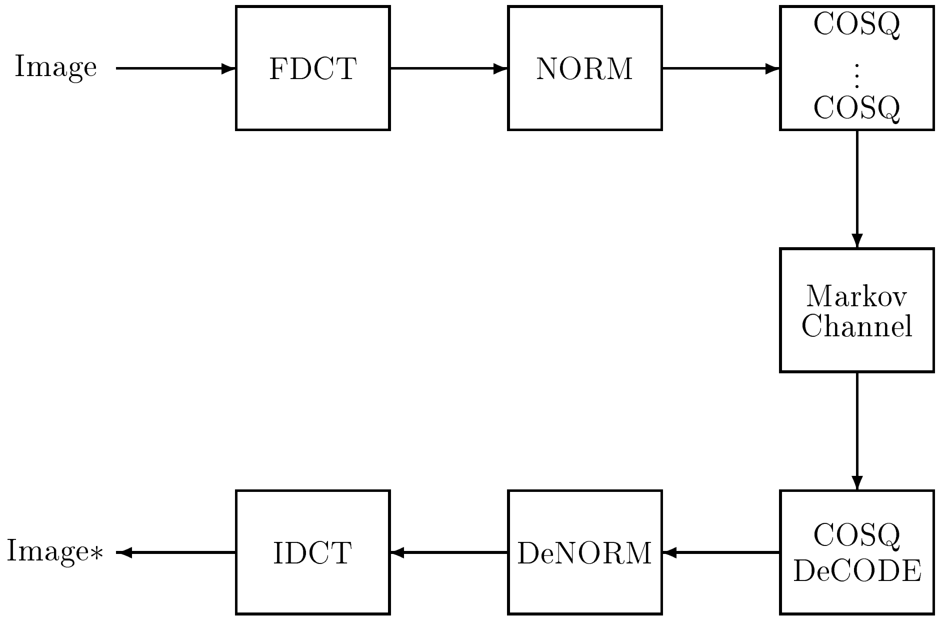


Figure 1: A DCT-COSQ image transmission system.

Compressed Lena at 1.19 bpp

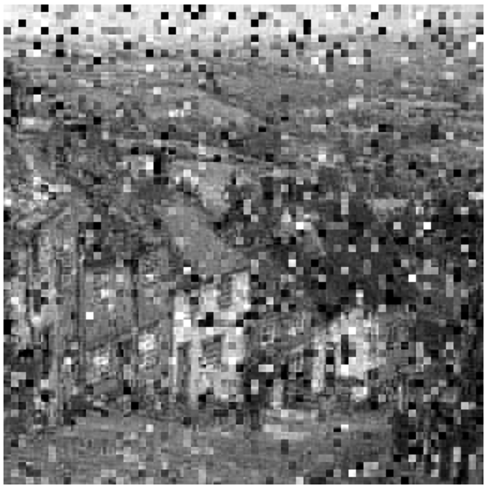
Decoded Lena with SQ-IL,
$PSNR = 15.84$ dB

Decoded Lena with CC-IL,
$PSNR = 17.05$ dB

Decoded Lena with COSQ,
$\delta = 10.0$, $PSNR = 27.88$ dB

Figure 2: Lena: Overall rate is 1.19 bpp; Markov channel with $\epsilon = 0.1$.

Compressed Goldhill at 0.90 bpp



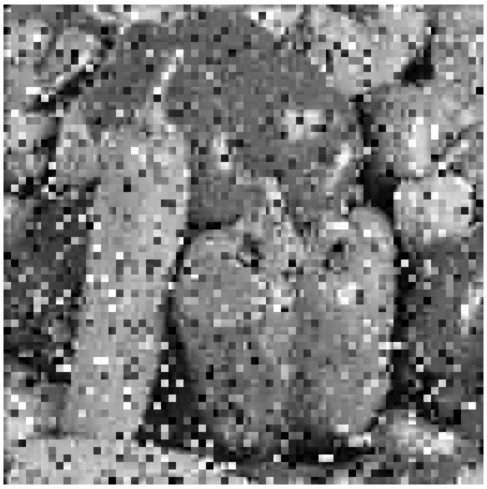
Decoded Goldhill with SQ-IL,
$PSNR = 16.32$ dB

Decoded Goldhill with CC-IL,
$PSNR = 17.87$ dB

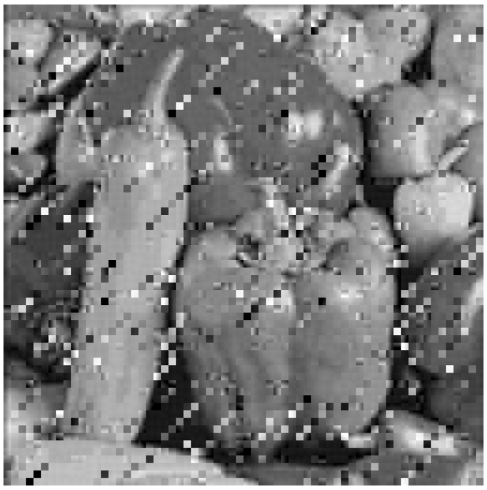
Decoded Goldhill with COSQ,
$\delta = 10.0$, $PSNR = 27.85$ dB

Figure 3: Goldhill: Overall rate is 0.90 bpp; Markov channel with $\epsilon = 0.1$.

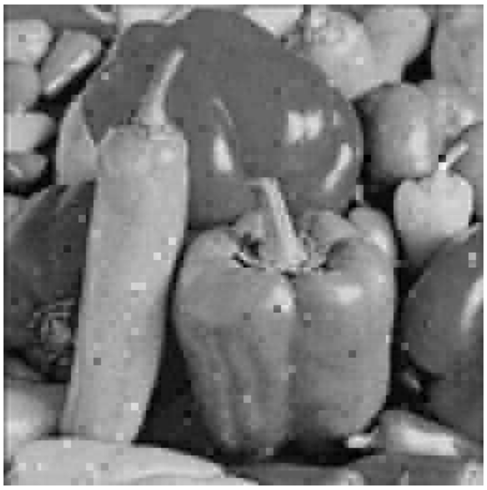
Compressed Peppers at 0.375 bpp



Decoded Peppers with SQ-IL,
$PSNR = 15.82$ dB



Decoded Peppers with CC-IL,
$PSNR = 17.82$ dB



Decoded Pepper with COSQ,
$\delta = 10.0$, $PSNR = 25.22$ dB

Figure 4: Peppers: Overall rate is 0.375 bpp; Markov channel with $\epsilon = 0.1$.

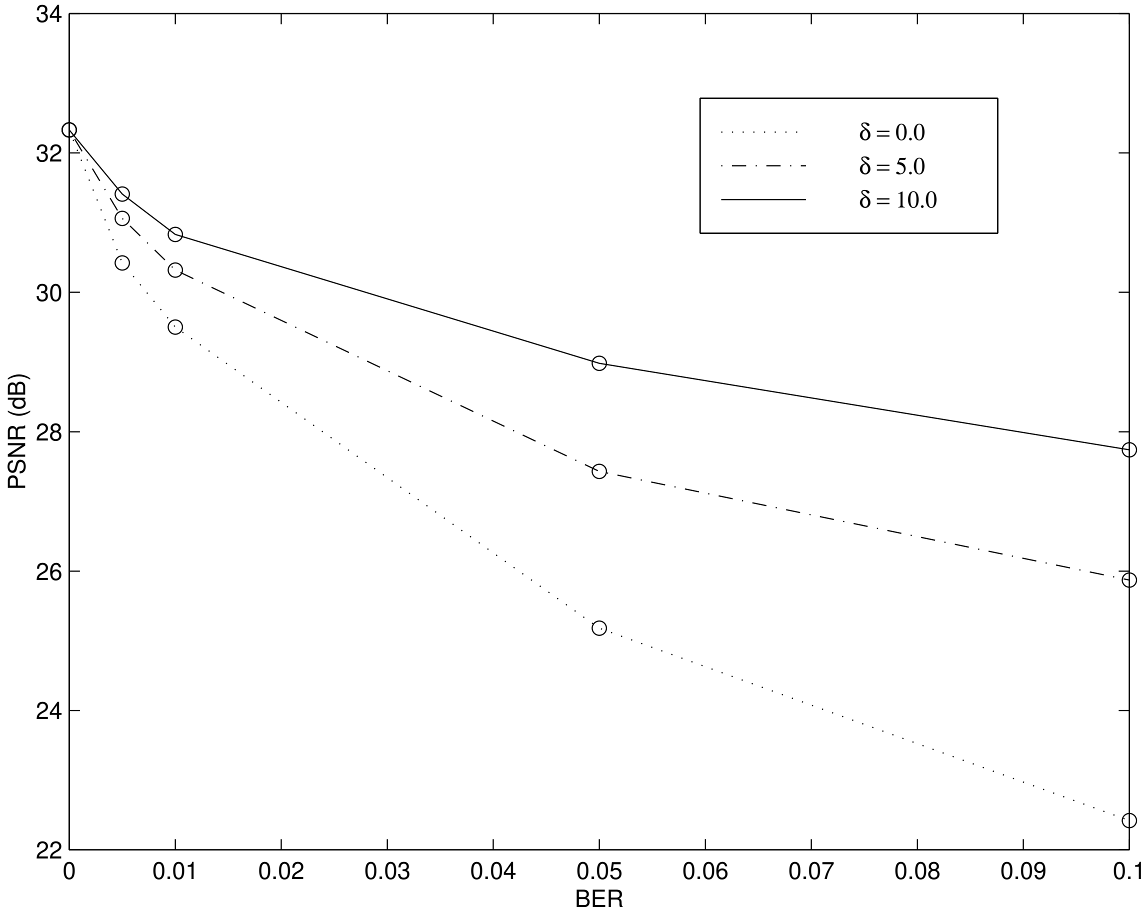


Figure 5: Lena: Gain due to memory over Markov channel; overall rate is 1.19 bpp.

δ	System	ϵ=0	ϵ=0.005	ϵ=0.01	ϵ=0.05	ϵ=0.1
0	COSQ	32.33	30.42	29.50	25.18	22.42
0	MAP-UNC [5]	31.75	26.29	23.93	17.76	15.25
0	MAP-UEP I [5]	31.75	29.70	28.37	23.17	18.28
5	COSQ	32.33	31.06	30.32	27.43	25.87
5	MAP-UNC [5]	31.75	25.59	23.28	18.38	16.38
5	MAP-UEP I [5]	31.75	30.77	29.93	24.49	18.83
10	COSQ	32.33	31.41	30.83	28.98	27.74
10	MAP-UNC [5]	31.75	26.20	24.07	19.04	17.17
10	MAP-UEP I [5]	31.75	31.02	30.32	24.77	18.95
0	SQ-IL	32.33	27.30	24.99	18.59	15.69
0	CC-IL	28.88	28.87	28.86	25.61	17.18

Table 1: Average PSNR (dB) of decoded Lena over the Markov channel ($M = 1$) with BER ϵ and correlation parameter δ using a fixed bit allocation table at 1.19 bpp. MAP-UEP I has a rate of 1.31 bpp [5].

δ	System	ϵ=0	ϵ=0.005	ϵ=0.01	ϵ=0.05	ϵ=0.1
0	COSQ	30.49	29.26	28.63	24.92	22.29
0	MAP-UEP II [5]	28.52	28.26	27.90	24.36	18.83
5	COSQ	30.49	29.58	29.01	26.93	25.58
5	MAP-UEP II [5]	28.52	28.39	28.20	24.66	18.80
10	COSQ	30.49	29.85	29.49	28.23	27.20
10	MAP-UEP II [5]	28.52	28.39	28.23	24.67	18.84
0	SQ-IL	30.49	26.57	24.56	18.50	15.68
0	CC-IL	28.28	28.27	28.26	25.33	17.19

Table 2: Average PSNR (dB) of decoded Lena over the Markov channel ($M = 1$) with BER ϵ and correlation parameter δ using a fixed bit allocation table at 0.90 bpp. MAP-UEP II hasa rate of 0.97 bpp [5].

δ	System	ϵ=0	ϵ=0.005	ϵ=0.01	ϵ=0.05	ϵ=0.1
0	COSQ	26.16	25.80	25.60	23.74	21.71
5	COSQ	26.16	25.91	25.07	24.72	23.98
10	COSQ	26.16	26.00	25.89	25.36	24.92
0	SQ-IL	26.16	24.37	23.06	18.19	15.58
0	CC-IL	25.53	25.52	25.51	23.88	17.48

Table 3: Average PSNR (dB) of decoded Lena over the Markov channel ($M = 1$) with BER ϵ and correlation parameter δ using a fixed bit allocation table at 0.375 bpp.

$B=76$								$B=58$								$B=24$							
8	7	6	4	3	0	0	0	8	7	6	4	0	0	0	0	8	8	0	0	0	0	0	0
7	6	5	4	0	0	0	0	7	6	5	0	0	0	0	0	8	0	0	0	0	0	0	0
6	5	4	0	0	0	0	0	6	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	4	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 4: Global fixed bit allocation tables for the DCT-COSQ system.

$B=76$								$B=58$								$B=24$							
7	6	5	1	0	0	0	0	7	6	3	0	0	0	0	0	4	4	0	0	0	0	0	0
6	5	1	0	0	0	0	0	6	4	0	0	0	0	0	0	4	0	0	0	0	0	0	0
5	1	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 5: Global fixed bit allocation tables for the CC-IL tandem system.

	$\delta = 0.0$	$\delta = 5.0$	$\delta = 10.0$
$\epsilon = 0.0$	$\begin{matrix} 7&6&4&4&3&2&1&0 \\ 5&4&4&3&2&2&0&0 \\ 3&3&3&3&2&1&0&0 \\ 2&2&2&2&2&0&0&0 \\ 1&1&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{matrix}$	$\begin{matrix} 7&6&4&4&3&2&1&0 \\ 5&4&4&3&2&2&0&0 \\ 3&3&3&3&2&1&0&0 \\ 2&2&2&2&2&0&0&0 \\ 1&1&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{matrix}$	$\begin{matrix} 7&6&4&4&3&2&1&0 \\ 5&4&4&3&2&2&0&0 \\ 3&3&3&3&2&1&0&0 \\ 2&2&2&2&2&0&0&0 \\ 1&1&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{matrix}$
$\epsilon = 0.005$	$\begin{matrix} 8&8&5&3&3&2&0&0 \\ 6&5&3&3&2&1&0&0 \\ 3&3&3&2&2&1&0&0 \\ 2&2&2&2&1&0&0&0 \\ 1&1&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{matrix}$	$\begin{matrix} 8&8&4&3&3&2&1&0 \\ 5&4&4&3&2&1&0&0 \\ 3&3&3&3&2&1&0&0 \\ 2&2&2&2&1&0&0&0 \\ 1&1&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{matrix}$	$\begin{matrix} 8&8&4&3&3&2&1&0 \\ 5&4&4&3&2&1&0&0 \\ 3&3&3&3&2&1&0&0 \\ 2&2&2&2&1&0&0&0 \\ 1&1&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{matrix}$
$\epsilon = 0.01$	$\begin{matrix} 8&8&5&3&2&2&0&0 \\ 7&5&3&3&2&1&0&0 \\ 3&3&3&2&2&1&0&0 \\ 2&2&2&2&1&0&0&0 \\ 1&1&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{matrix}$	$\begin{matrix} 8&8&4&3&3&2&0&0 \\ 6&4&4&3&2&1&0&0 \\ 3&3&3&3&2&1&0&0 \\ 2&2&2&2&1&0&0&0 \\ 1&1&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{matrix}$	$\begin{matrix} 8&8&5&3&3&2&0&0 \\ 5&5&3&3&2&1&0&0 \\ 3&3&3&3&2&1&0&0 \\ 2&2&2&2&1&0&0&0 \\ 1&1&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{matrix}$
$\epsilon = 0.1$	$\begin{matrix} 8&8&8&4&2&1&0&0 \\ 8&8&6&4&1&0&0&0 \\ 4&4&4&1&1&0&0&0 \\ 1&1&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{matrix}$	$\begin{matrix} 8&8&7&4&2&1&0&0 \\ 8&7&5&3&2&0&0&0 \\ 4&4&3&2&1&0&0&0 \\ 2&2&2&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{matrix}$	$\begin{matrix} 8&8&8&3&3&2&0&0 \\ 8&6&5&3&2&0&0&0 \\ 3&3&3&3&2&0&0&0 \\ 2&2&2&2&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \end{matrix}$

Table 6: Optimal bit allocation matrices for Lena at B=76 bits, or 1.19 bpp.

δ	System	ϵ=0	ϵ=0.005	ϵ=0.01	ϵ=0.05	ϵ=0.1
0	COSQ-OPT	33.07	30.71	29.74	25.43	22.55
0	COSQ-FIXED	32.33	30.42	29.50	25.18	22.42
5	COSQ-OPT	33.07	31.48	30.61	27.55	26.00
5	COSQ-FIXED	32.33	31.06	30.32	27.43	25.87
10	COSQ-OPT	33.07	31.88	31.32	29.17	28.11
10	COSQ-FIXED	32.33	31.44	30.83	28.98	27.74

Table 7: Performance comparison between COSQ systems using fixed and optimal bit allocation tables; PSNR (dB) of decoded Lena over the Markov channel with BER ϵ and correlation parameter δ at 1.19 bpp.

δ	System	ϵ=0	ϵ=0.005	ϵ=0.01	ϵ=0.05	ϵ=0.1
0	COSQ-OPT	31.76	29.87	29.05	25.07	22.40
0	COSQ-FIXED	30.49	29.26	28.63	24.92	22.29
5	COSQ-OPT	31.76	30.40	29.62	27.12	25.75
5	COSQ-FIXED	30.49	29.58	29.01	26.93	25.58
10	COSQ-OPT	31.76	30.83	30.25	28.65	27.43
10	COSQ-FIXED	30.49	29.85	29.49	28.23	27.20

Table 8: Performance comparison between COSQ systems using fixed and optimal bit allocation tables; PSNR (dB) of decoded Lena over the Markov channel with BER ϵ and correlation parameter δ at 0.9 bpp.

δ	System	ϵ=0	ϵ=0.005	ϵ=0.01	ϵ=0.05	ϵ=0.1
0	COSQ-OPT	27.91	26.85	26.41	23.81	21.74
0	COSQ-FIXED	26.16	25.80	25.60	24.74	21.71
5	COSQ-OPT	27.91	27.00	26.76	25.24	24.29
5	COSQ-FIXED	26.16	25.91	25.70	24.72	23.98
10	COSQ-OPT	27.91	27.44	27.14	26.22	25.39
10	COSQ-FIXED	26.16	26.00	25.89	25.36	24.92

Table 9: Performance comparison between COSQ systems using fixed and optimal bit allocation tables; PSNR (dB) of decoded Lena over the Markov channel with BER ϵ and correlation parameter δ at 0.375 bpp.

	ϵ_d=0.0	ϵ_d=0.01	$\epsilon_d = 0.1$
ϵ_a=0.0	32.33	31.88	31.02
ϵ_a=0.01	24.97	30.83	30.52
ϵ_a=0.1	15.63	26.28	27.74

Table 10: Performance (in dB) of decoded Lena under mismatch in ϵ ($\delta_a = \delta_d = 10.0$).

	δ_d=0.0	δ_d=5.0	$\delta_d = 10.0$
δ_a=0.0	29.50	26.29	24.85
δ_a=5.0	26.33	30.32	30.01
δ_a=10.0	26.66	30.82	30.83

Table 11: Performance (in dB) of decoded Lena under mismatch in δ ($\epsilon_a = \epsilon_d = 0.01$).

δ	System	ϵ=0	ϵ=0.005	ϵ=0.01	ϵ=0.05	ϵ=0.1
0	COVQ	30.07	29.27	28.37	24.43	21.83
0	VQ-IL	30.07	22.24	19.67	13.56	11.23
5	COVQ	30.07	30.01	29.30	26.49	25.07
5	VQ-IL	30.07	22.24	19.67	13.56	11.23
10	COVQ	30.07	30.04	29.73	27.84	26.54
10	VQ-IL	30.07	22.24	19.67	13.56	11.23

Table 12: Average PSNR (dB) of decoded Lena over the Markov channel ($M = 1$) using COVQ with dimension $= 4 \times 2$, $R = 1$ bit per sample.