Multimode Image Coding for Noisy Channels

Shankar L. Regunathan, Kenneth Rose and Shrinivas Gadkari
Department of Electrical and Computer Engineering
University of California
Santa Barbara, CA 93106

Abstract— **We attack the problem of designing robust image compression schemes for transmission over noisy channels. To achieve the dual goals of high compression efficiency and low sensitivity to channel noise, we introduce a multimode coding framework. Multimode coders are quasi-fixed length in nature, and this enables optimal tradeoff between the compression capability of variable-length coding and the robustness of fixed length coding. The framework lends itself to a natural incorporation of conventional source-channel coding mechanisms, and thus provides a joint design methodology for the source and channel coding components. We first derive a general Lagrangian based algorithm to design the multimode coder which optimizes the overall rate-distortion performance. We then apply our framework to develop multimode image coding (MIC) schemes for noisy channels. We demonstrate the benefits of the multimode approach by developing two specific image compression algorithms in conjunction with the following source-channel coding mechanisms: (i) channel optimized quantizers, and (ii) transmission energy allocation. Simulations demonstrate that substantial gains of up to 7dB can be achieved by MICs over conventional approaches.**

I. Introduction

Although much of the image coding literature has ignored issues of transmission over noisy channels, the topic has recently gained in importance due to emerging "hot" applications such as multimedia communications over wireless channels. Moreover, separate handling of source and channel coding, while asymptotically justifiable by Shannon's theory, may be significantly suboptimal in practice given the demanding requirements on bit rate, complexity, delay, and error robustness to noisy time-varying channels. Thus, several researchers have investigated the applicability of combined source-channel coding to robust image compression leading to notable improvements over unprotected image coding [1], [2], [3], [4].

A major shortcoming of the standard source-channel image coding techniques is due to their use of fixed-length encoding. It is well known that, as long as vector dimensions are not exceedingly high, fixed length source coders are significantly inferior to variable length coders in their compression efficiency. More importantly, images are well modelled as a mixture of multiple sources [7], [8] and a variable length coder can be tailored to exploit these highly non-stationary characteristics, and provide additional gains over fixed-length coding.

However, conventional variable length codes are extremely sensitive to channel errors that may cause desynchronization and catastrophic error propagation. In contrast, fixed length coding offers greater robustness to channel errors, as an error is confined to just one codeword. Thus, it is not surprising that image coders for noisy channel environments predominantly use fixed-length coding (e.g., [1],[2],[3]), and their compression performance is sacrificed for error resilience.

Ideally, we would like to achieve the best of both worlds, namely, robustness to channel errors and high compression efficiency. We propose a method to optimize the tradeoff between the two competing objectives via the framework of *multimode coding*. The multimode coder switches between different fixed length codes for each block of data to achieve efficient compression. Yet, as long as the mode information is uncorrupted by noise (which is ensured through heavy protection), there is no error propagation in the fixed length part. We exploit this *quasi-fixed length* nature of multi-mode coding in designing robust image compression schemes which substantially outperform conventional approaches.

The rest of the paper is organized as follows: The generic multimode coding framework for noisy channels is introduced in section II. We discuss its advantages over conventional coding schemes and propose a simple design algorithm which jointly optimizes compression efficiency and robustness by iteratively minimizing the rate-distortion cost. In section III, we apply this framework to the design of efficient multimode coding schemes using the adaptive discrete cosine transform as the starting point. Specific MIC algorithms are developed in conjunction with the following source-channel coding methods: channel optimized quantizers [4], and the recent method of transmission energy allocation [6]. Section IV presents the simulation results which demonstrate the improvements in performance achievable by the multimode coding approach over conventional approaches.

II. Multimode Coding

The basic idea in multimode coding is to allow a set of possible modes in which the coder can operate, where each mode is, in fact, a fixed rate encoding algorithm. For each block of data, the encoder can choose the best mode for operation given the local statistics (or other parameters) so as to achieve the optimum overall rate-distortion performance. The encoding mode is heavily protected and sent as side information.

A. Structure

Let $\{X\}$ represent k-dimensional random vectors generated from a source $\mathcal{X}$. Let $\{m_j\}$, $j = 1, 2, ...M$ denote the available modes. Let each mode m_j be associated with its own fixed length encoding (and decoding) scheme whose rate in bits per source vector is denoted by r_j.

For each source vector, the encoder selects a particular mode m_j and uses the associated encoding scheme. The mode information is transmitted to the decoder as heavily protected side information. We will assume, for the time being, that probability of the error in the mode information is negligible due to this protection. Let c_j denote the total rate (including protection) for specifying to the decoder that mode m_j is used. The total rate needed for encoding some source vector x using the mode m_j is thus

$$R(x) = r_j + c_j \tag{1}$$

Since the decoder has perfect information about the mode m_j, it uses the corresponding decoding algorithm to produce an estimate $\hat{x}$ which in general is corrupted both by compression and channel noise. The expected rate R for encoding the source is then,

$$R = E(R(X)) \tag{2}$$

where the expectation is over the source statistics. The average distortion D is give by,

$$D = E(d(X, \hat{X})) \tag{3}$$

where the expectation is over both the source and channel statistics, and $d(\cdot, \cdot)$ is a suitably defined distortion measure.

The objective is to minimize the distortion D while satisfying the constraint on the rate R. Since D represents the overall reconstruction error (due to both source compression and channel noise) and R is the total rate, we are indeed optimizing the overall rate-distortion cost. In subsection II-C, we describe a design algorithm to achieve this objective.

B. Advantages

The multimode coding structure provides practical means to achieve both compression efficiency and error resilience. Adaptation to varying source statistics is achieved by switching the encoding mode. Thus, to a certain degree, multimode codes retain the flexibility of variable length codes which enables efficient compression of sources with nonstationary statistics such as images. On the other hand, the heavy protection of mode information ensures its uncorrupted transmission, so that synchronization is always maintained at the decoder. Therefore, the data is effectively transmitted in a fixed length manner. This quasi-fixed length operation eliminates significant propagation of channel error. Note that by careful design, the mode information can be made a very small part of the total rate. Thus the overall compression performance is not significantly impaired by the heavy protection of mode information.

The number of available modes (the mode information rate) determines the flexibility of the coder. On one extreme, a large number of modes ensures that the code is well adapted to a variety of statistics. However, the increase in overhead information must be taken into account particularly due to its heavy protection. A completely fixed length coder is the other extreme where there is only one mode. Here, no mode information needs to be sent and correspondingly there is no flexibility in adaptation. In multimode coders, we achieve efficient rate-distortion performance by optimizing this tradeoff.

Moreover, within each mode, appropriate fixed length codes can be chosen from the available conventional (fixed length) source-channel coding mechanisms and this integration leads to a complete multimode coder. The objective of the proposed algorithm is to jointly design the modes as well as the fixed length codes of the multimode coder by optimizing the rate-distortion cost.

C. Optimization

A training set $\{x_i\}$, $i = 1, 2, ...N$ is generated from the source $\mathcal{X}$. Replacing the expectation over the source statistics by sample average over the training set, the objective becomes one of minimizing the distortion (where the expectation is only over the channel statistics)

$$D = \frac{1}{N} \sum_{i=1}^{N} E(d(x_i, \hat{x}_i)) \tag{4}$$

subject to the rate constraint,

$$R = \frac{1}{N} \sum_{i=1}^{N} R(x_i) \leq R_{max} \tag{5}$$

We naturally rewrite this constrained optimization problem as an unconstrained minimization of the Lagrangian $L = D + \lambda R$, where λ is the Lagrange multiplier. The multimode coder can also be regarded as a two stage code albeit with heavily protected first stage. From this viewpoint, we propose an iterative algorithm similar to [8] to solve this design problem. Note that the objective in [8] was one of pure source coding while our objective is one of robust coding for transmission through noisy channels.

Algorithm: Partition the training set into the M modes. This initial partition could be arbitrary or based on some "smart" heuristic. Iterate the following steps:

1. For $j = 1, 2, ..M$, optimize the fixed length code to minimize the Lagrangian for the training subset of mode m_j.
2. Design optimal indices (with protection) to represent each mode, based on the size of the training subsets, so as to minimize the average side information rate.
3. Repartition the training set. i.e., assign each training vector to the (optimal) mode that minimize its contribution to the Lagrangian L.

Each step of the algorithm is non increasing in the Lagrangian cost and thus the design method results in a locally optimal multimode coding scheme.

III. Multimode Image Coding (MIC)

The adaptive discrete cosine transform (ADCT) [9] is a successful method for image coding targetted at noiseless channels. We use ADCT as a starting point to develop multimode image coding (MIC) schemes which provide efficient compression as well as robustness to channel errors.

The image is divided into a set of disjoint blocks $i = 1, 2, ...N$. A two dimensional DCT is applied to each block. Each DCT block is classified into one of M modes, where each mode is associated with a bit allocation map. The DCT coefficients are scalar quantized according the number of bits specified by the map and transmitted. (Note that, in principle, any fixed length vector quantization scheme of the block can replace scalar quantization.) The mode index is transmitted to the decoder as side information. It is protected by a rate 1/3 error correction code which ensures that the mode information is transmitted reliably with negligible error probability [2]. The decoder uses the received values and the associated mode index to produce an estimate of the original DCT block. Next, it performs the inverse DCT to reconstruct the image. The unitary nature of the transform ensures that the image reconstruction error (measured by the squared error criterion) is equal to the error in the transform domain.

A complete MIC is designed by the integration of a source-channel coding scheme (as the fixed length code), suited for the given complexity constraints and channel conditions. We choose to demonstrate the power of multimode image coding by developing specific schemes in conjunction with the following techniques: (i) channel-optimized quantizers (COQ) [5], and (ii) the recently developed method of transmission energy allocation [6].

A. MIC-COQ: with Channel Optimized Quantizers

Channel optimized quantizer (COQ) is a convenient source-channel coding method where the quantizer performs both source compression and channel protection. COQ is designed by a modified version of the generalized Lloyd algorithm which takes into account the effect of channel errors (see e.g. [4]). We next describe the design of a complete multimode image coding scheme incorporating COQs.

Let ϵ be the bit error rate on the given binary symmetric channel. The design objective is to optimize the bit allocation maps, mode indices and channel optimized quantizers. To simplify the quantizer design, we assume that the probability densities of DCT coefficients can be reasonably approximated by gaussian distributions [7]. We also impose the requirement that no coefficient be encoded using more than r_{max} bits. We design a set of quantizers of rates $r = 1, 2, ..., r_{max}$ bits, where each quantizer is optimized for a unit variance Gaussian source and the given channel. Let us denote these quantizers by $\{Q^0(1), Q^0(2), ..., Q^0(r_{max}\}$ and

the corresponding distortion they produce by $\{d(r)\}$. The channel-matched quantizer for a Gaussian variable of variance σ^2 with rate of r bits, is obtained through scaling of $Q^0(r)$ by a factor of σ; the resulting distortion is given by $\sigma^2\, d(r)$.

We start with an initial partition of the training set based on the AC energy of the blocks. Let the total number of DCT coefficients in each block be k. The corresponding steps in the iterative design algorithm are:

1. Given the current training set partition, design the quantizers and bit allocation maps for each class: For each coefficient $i = 1, 2, ..., k$, and for each class $j = 1, 2, ..., M$:
   (a) Compute σ_{ij}^2, the empirical variance of the ith coefficient of blocks belonging to mode j.
   (b) Optimize bit allocation as $r_{ij} = \arg\min_r\{\sigma_{ij}^2 d(r) + \lambda r\}$.
   (c) The quantizer Q_{ij} is obtained by scaling $Q^0(r_{ij})$ by σ_{ij}.
2. Redesign the mode indices based on the empirical rate of occurrence of each mode.
3. Given the new quantizers and prefix code, repartition the training data.

B. MIC-TEA: Multimode Image Coding with Transmission Energy Allocation

Transmission energy allocation (TEA) [6] is a simple and robust protection mechanism for time varying channels such as wireless channels. Traditional combined source-channel coding techniques are optimized for a fixed known channel statistics and suffer degradation under conditions of channel mismatch. TEA combines source-channel coding with modulation optimization, and has been shown to achieve good performance for a wide range of channel conditions.

TEA consists of the following steps: i) Design the quantizers for the noiseless channel so as to retain optimal performance when the channel is clean; ii) use natural binary code to achieve an hierarchical ordering of bits; and iii) provide a varying level of protection to bits (according to their importance) by optimizing of the allocation of transmission energy to them. As the channel gets noisier, we lose the less significant information first, thus resulting in a steady but graceful degradation in performance.

We first give a review of the TEA algorithm in a general setup [6]. Consider a n-bit quantizer designed for noiseless channel and indexed by a natural binary code. Let its codebook $C = \{y_0, y_1, ..., y_{2^n-1}\}$. Let the quantizer encode an input vector x to y_i. Let the decoded output be y_j which may be different from y_i due to channel noise. The average distortion is given by

$$D = E||x - y_j||^2 = E||x - y_i||^2 + E||y_i - y_j||^2 \quad (6)$$

since $\{y_i\}$ are at the centroids of their encoding regions. The first term corresponds to the distortion of the source encoder which is fixed for a given source/quantizer pair and independent of the channel. The second term can be interpreted as the channel distortion(D_c).

We assume a transmission scheme where each of the n bits in the codeword is transmitted independently. Let E_{tot}

denote the total quota of transmission energy available for allocation among the n bits. Let E_i denote the energy allocated to the ith bit. The probability of channel error for the ith bit can be written as $\epsilon_i = f(E_i, \sigma_n^2)$, where σ_n^2 is the variance of the representative level of Gaussian channel noise, and the function $f(\cdot)$ depends on the type of modulation used in the system.

To allocate the transmission energy, we first evaluate the error sensitivity of the different bits in the codeword. The sensitivity S_i of the ith bit is defined as the expected amount of distortion caused by a bit error at this location. If the bit error rate is sufficiently small, the probability of more than a single bit error is negligible and the channel distortion D_c can be expressed as,

$$D_c = \sum_{i=1}^{n} \epsilon_i S_i \tag{7}$$

We allocate the transmission energy so as to minimize the Lagrangian contribution of the channel distortion, subject to the total energy constraint:

$$\sum_{l=1}^{n} E_l = E_{tot}.$$

The necessary condition for optimality is:

$$\frac{\partial D_c}{\partial E_i} = D_i \frac{\partial \epsilon_i}{\partial E_i} = -\lambda, \qquad \text{if } E_i > 0 \tag{8}$$

A greedy allocation algorithm, similar to standard bit allocation algorithm, is described in [6]. Note that no additional redundancy is introduced by energy allocation thus retaining the performance for error free channel.

We next describe the incorporation of TEA in the multimode image coding system. We follow the same design steps described in section II-C but optimize the bit allocations and the quantizers for the *error free channel*. The quantizers are indexed by a natural binary code.

For simplicity, assume BPSK as the modulation scheme. Let the system be allowed an energy E_b per bit and let be σ_n^2 be the variance of the channel noise . If energy allocation was allocated uniformly to all the bits (as is commonly done), the channel would be equivalent to a BSC channel of transition error probability $\epsilon = f(E_b, \sigma_n^2)$. We perform TEA on each mode of the coding scheme separately. Let R_k be the total rate of mode k. Hence, the total energy available to us for allocation among the R_k bits is $E_{tot} = R_k E_b$. We evaluate the sensitivity of each of these bits using (6) and apply TEA to distribute total energy E_{tot}, optimally to the R_k bits.

IV. Results

We now present the simulation results obtained by using multimode coders to compress real world images and transmitting through noisy channels. The training set was generated from the image "BARBARA" and used to design multimode coding algorithms MIC-COQ and MIC-TEA , with the number of modes $N = 1, 4$ and 16 in both the cases. Table 1 lists the PSNR values achieved on the test set image "LENA" and table 2 on the image "PEPPER", for rate R (in bpp) in the range of 0.4 to 1.0 and a channel transition error probability $\epsilon = 0.005$. The performance of MIC under conditions of channel mismatch is shown in table 3. The listed rates include the rate required for transmitting the protected mode information. Note that the single mode ($N = 1$) cases in MIC-COQ corresponds to the fixed rate coder of [2]. In all cases, the side information (if any) was protected by a rate 1/3 convolutional code and was assumed to be transmitted error free [4]. It can be seen that the proposed multi-mode coding outperforms fixed rate coding, and achieves dramatic performance gains of up to 6 dB. Note also that the gains increase with bit rate since the mode information becomes a smaller fraction of the overall bit rate.

R		0.4	0.5	0.6	0.75	1.0
	N					
MIC COQ	1	24.15	24.19	24.80	25.37	26.67
	4	28.82	29.54	30.12	30.82	31.67
	16	29.54	30.54	31.34	32.17	32.96
MIC TEA	1	24.19	24.82	24.99	25.61	26.88
	4	29.01	29.82	30.43	31.94	33.19
	16	30.35	31.24	33.00	33.71	34.43

TABLE I

Performance of multimode coding (PSNR in dB) on test image "LENA" at various rates R (measured in bpp) and channel bit error rate $\epsilon = 0.005$. $N > 1$ is multimode coding while $N = 1$ corresponds to Fixed Rate Coding as in [1].

R		0.4	0.5	0.6	0.75	1.0
	N					
MIC COQ	1	22.20	22.29	22.75	23.13	24.04
	4	26.47	26.91	27.20	27.75	28.36
	16	27.50	28.31	28.75	29.39	30.02
MIC TEA	1	22.44	22.77	22.86	23.31	24.12
	4	26.27	26.99	27.89	28.22	29.39
	16	27.89	28.30	29.60	29.95	30.15

TABLE II

Performance of multimode coding (PSNR in dB) on test image "PEPPER" at various rates R (measured in bpp) and channel bit error rate $\epsilon = 0.005$. $N > 1$ is multimode coding while $N = 1$ corresponds to Fixed Rate Coding as in [1].

V. Conclusion

We have proposed a new multimode framework for robust image coding. The framework is quasi-fixed in nature and allows explict tradeoff between the compression capability of variable length coding and robustness of fixed length coding. We proposed a Lagrangian-based design algorithm for multimode coders which directly optimizes

ϵ_a		0.0	0.005	0.01
	N			
MIC COQ	1	27.20	26.67	26.02
	4	32.87	31.67	30.20
	16	34.32	32.96	31.24
MIC TEA	1	28.02	26.88	25.92
	4	35.21	33.19	31.25
	16	37.20	34.43	32.37

TABLE III

PERFORMANCE OF MIC ON TEST IMAGE "LENA" UNDER CHANNEL MISMATCH CONDITIONS. $N = 1$ IS FIXED LENGTH CODING AND $N > 1$ CORRESPONDS TO MULTIMODE CODING. THE CODERS WERE DESIGNED FOR CHANNEL ERROR RATE ϵ_d =0.005 AND TESTED UNDER CHANNEL ERROR RATES ϵ_a = 0.0, 0.005 AND 0.01. THE BIT RATE R = 1.0 BPP.

the rate-distortion performance. The framework allows the incorporation of various channel protection methods (as suitable fixed length codes) to enhance the performance of the coding system. We then developed specific multimode image coding schemes (MIC) in conjunction with channel optimized quantizers (COQ), and transmission energy allocation (TEA). Simulation results demonstrate that significant gains of upto 7 dB can be achieved by multimode image coding (MIC). Multimode image coding with transmission energy allocation MIC-TEA, is a very suitable candidate for image compression aimed at transmission over noisy time varying channels.

Current work involves applying this framework to subband image coding for noisy channels, and results will be presented in a future publication.

Fig. 1. Performance of MIC-COQ on test image "LENA" for N = 1(top), and 16 (bottom) modes. The bit rate R = 1.0 bpp and channel error rate ϵ = 0.005

REFERENCES

[1] J. W. Modestino, D. G. Daut and A. L. Vickers, "Combined source-channel coding of images using the block cosine transform," *IEEE Trans. on Comm.*, vol. 29, pp. 1261-74, Sep. 1981.

[2] V. Vaishampayan and N. Farvardin, "Optimal block cosine transform image coding for noisy channels," *IEEE Trans. on Comm.*, Mar. 1990, vol. 38, pp 327-36.

[3] M. J. Ruf, "A high performance fixed rate compression scheme for still image transmission," *Proc. DCC*, 1994, pp. 294-303.

[4] N. Tanabe and N. Farvardin, "Subband image coding using entropy-coded quantization over noisy channels," *IEEE JSAC.*, Jun. 1992, vol. 10, pp. 926-42.

[5] N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channel: An approach to combined source-channel coding," *IEEE Trans. on Info. Theory*, Nov. 1987, vol. IT-33, pp 827-38.

[6] S. Gadkari and K. Rose, "Robust Vector Quantization by Transmission Energy Allocation," *Electronics Letter*, Aug. 1996, vol. 32, no. 16, pp. 1451-53.

[7] J. M. Lervik and T. A. Ramstad, "Optimality of multiple entropy coder systems for nonstationary sources modelled by a mixture distribution," *Proc. ICASSP*, 1996.

[8] M. Effros and P. A. Chou, "Weighted Universal Bit Allocation: Optimal multiple quantization matrix coding," *Proc. ICASSP*, 1995, pp. 2343-46.

[9] W. H. Chen and C. H. Smith, "Adaptive Coding of Monochrome and Color images,"*IEEE Trans. on Comm.*, Vol. COM-25, pp. 1285-1292, Nov. 1977.