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 robust and efficient image compression 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 allow optimization of the tradeoff between the compression capability of variable-length coding and the robustness to channel errors of fixed length coding. We apply our framework to develop multimode image coding (MIC) schemes for noisy channels, based on the adaptive DCT. The robustness of the proposed MIC is further enhanced by the incorporation of a channel protection scheme suitable for the constraints on complexity and delay. To demonstrate the power of the technique we develop two specific image coding algorithms optimized for the binary symmetric channel. The first, MIC1, incorporates channel optimized quantizers and the second, MIC2, uses rate compatible punctured convolutional codes within the multimode framework. Simulations demonstrate that the multimode coders obtain significant performance gains of up to 6dB over conventional fixed length coding techniques.

1 Introduction

Although much of the image coding literature has ignored issues of transmission errors due to noisy channels, the topic has been gaining in urgency due to emerging "hot" applications such as multimedia communications over wireless channels. Moreover, separate handling of source and channel coding, while asymptotically justifiable by

*This work was supported in part by by National Science Foundation under grant no. NCR-9314335, University of California MICRO program, ACT Networks, Advanced Computer Communications, Stratacom, DSP Group, DSP Software Engineering, Fujitsu, General Electric Company, Hughes Electronics, Intel, Moseley Associates, National Semiconductor, Nokia Mobile Phones, Qualcomm, Rockwell International, and Texas Instruments.

Shannon's theory, appears too inefficient in practice given the demanding requirements on bit rate, complexity, delay, and of robustness to noisy time-varying channels. Thus, several researchers have been investigating the applicability of combined source-channel coding to robust image compression [1], [2], [3].

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 excessively high, fixed length source coders are significantly inferior to variable length coders in terms of compression efficiency. More importantly, images are well modelled as a mixture of multiple sources [9], [10] and a variable length coding scheme can be tailored to exploit these highly non-stationary statistics, 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. Fixed length coding offers greater robustness to channel errors, as an error is confined to just one codeword. Indeed, image coders for noisy channel environments predominantly use fixed-length coding (e.g., [1], [2] , [3]), and 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. In this paper, 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. Moreover, as long as the mode information is error free, 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 other standard approaches.

The rest of the paper is organized as follows: We introduce the general multimode coding structure in section 2. We discuss its advantages over standard coding schemes and propose a simple optimization scheme for the design of the multimode coder. In section 3, the adaptive discrete cosine transform (ADCT) is viewed as a special case of multimode coding. We use this observation as the starting point for developing ADCT-based multimode image coding (MIC) schemes. In section 4, we incorporate two channel protection methods within our ADCT-based MIC framework. These are: channel optimized quantizers (e.g. [4]) and rate compatible punctured convolutional codes [5]. To demonstrate the performance gains achievable by the multimode coding approach over conventional image coding schemes, we provide simulation results for transmission over a binary symmetric channel.

2 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 mode, as well as the encoded data, are transmitted through a noisy channel.

2.1 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/decoding scheme whose rate in bits per source vector is denoted by r_j.

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

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

The transmitted value of X is corrupted by channel noise. Since the decoder has perfect information about the mode m_j, it can use the corresponding decoding algorithm to produce an estimate $\hat{X}$. The expected rate R for encoding the source is

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

where the expectation is over the source statistics. The expected distortion D is

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

where the expectation is over both the source and channel statistics and $d(\cdot, \cdot)$ is a suitably defined distortion measure. The design objective is to minimize the distortion D while satisfying the constraint on the rate R. In section 2.3, we describe a design method for this system.

2.2 Motivation

The quasi-fixed length operation of the multimode coder can achieve both compression efficiency and error resilience. The coder adapts its operation to the source by switching modes. Thus we retain the flexibility of variable rate coding which enables efficient compression of non stationary sources (images). Moreover, the heavy protection of the mode information ensures that it is exactly known at the decoder, and hence the actual data is effectively transmitted in a fixed length manner without significant error propagation. Note that by careful design, the mode information can be made a very small part of the total rate so that protecting the mode heavily does not impair the overall compression performance significantly. Moreover, within each mode, the effects of channel errors on the quantized data can be further reduced by incorporating standard channel protection techniques.

We note that the number of modes (and the mode information rate) determines the number of different fixed length codes (flexibility of the coder). On one extreme, a large number of modes ensures that the coder can be extremely adaptive to a variety of statistics. However, the overhead in mode information is greatly increased particularly due to its heavy protection. A completely fixed length coder is the other extreme where there is only one mode (no mode information needed) and correspondingly no flexibility in adaptation. We develop an optimization procedure to tradeoff the dual objectives of compression efficiency and robustness.

2.3 Optimization

A training set $\{x_i\}$, $i = 1, 2, ...N$ is generated from the source $\mathcal{X}$. Replacing the expectation over the source statistics with sample average over the training set, the objective of the coder becomes that of minimizing

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

subject to

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

where the expectation is now only over the channel statistics.

We naturally rewrite this constrained optimization problem as minimization of the Lagrangian $L = D + \lambda R$, where λ is the Lagrange multiplier. The resulting unconstrained minimization problem is separable, and the Lagrangian contribution of each training vector can be minimized independently. Note that the multimode coder can also be regarded as a two stage coder albeit with a heavy protection to the first stage. From this viewpoint, we adopt an iterative algorithm from [10] to design the multimode coder. It should however be emphasized that the objective in [10] 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 $j = 1, 2, ..M$, design an optimal fixed length code to minimize the lagrangian for the training subset of mode m_j.
2. Design optimal code words (with protection) c_j to represent each mode based on the population of training subsets to minimize the average side information rate.
3. Repartition the training set into modes such that each vector is encoded by its optimal mode. i.e., assign each training vector to the mode that minimizes its contribution to the Lagrangian L.

Each step of the algorithm is non decreasing in the Lagrangian cost and thus the design method produces a locally optimal multimode coding scheme for noisy channels. Note that the multimode coding framework is general and can encompass a wide variety of coding schemes. For example, most variable rate speech coding methods are in fact special cases of multimode coding, where the modes are "voiced frame", "unvoiced frame", "silence", etc. In this work, however, we restrict ourselves to applying this framework to design efficient and robust multimode image coders for noisy channels.

3 ADCT-based Multimode Image Coding (MIC)

An important special case of multimode image coding is the method of adaptive DCT (ADCT) [6]. The 2-dimensional separable DCT has been successfully employed as a decorrelating transform for image coding. Moreover, the DCT is image independent and can be efficiently implemented using fast algorithms. ADCT is a technique that exploits these features of DCT, while allowing flexibility in locally adapting the bit rate, and most importantly for our purposes, can be extended to provide robustness to channel errors. We now develop ADCT-based multimode image coding (MIC) schemes.

In the ADCT method, the image is divided into a set of disjoint blocks. The two dimensional DCT is applied to each block individually to get a vector of DCT coefficients $\{x_i\}, i = 1, 2, ..N$. Let the total number of DCT coefficients in each block be k. Each DCT block is classified into a mode $\{m_j\}$, $j = 1, 2, ...M$. The mode index $\{c_j\}$ is transmitted to the decoder as side information. We use a rate $1/3$ error correction code for protection which ensures that the mode information is transmitted reliably with a negligible rate of errors [8]. Each DCT coefficient is scalar quantized separately using the number of bits specified by the bit allocation map associated with the mode index. The decoder uses the received values and the associated mode index to make an estimate $\{\hat{x}_i\}$ of the DCT coefficients. Inverse DCT is then performed to reconstruct the image blocks. For the mean squared error distortion measure, the error in reproducing the image is equal to the error in reproducing the DCT coefficients since the DCT is a unitary transform.

Based on a training set of DCT blocks, we can apply the iterative algorithm to design bit allocations, fixed length codes for each coefficient and codewords to represent the modes.

4 Channel Protection in MIC

We can enhance the robustness of the scheme by incorporating channel protection schemes which are suited for the various constraints on complexity and delay. Examples include index assignment [12] and the recently developed method of transmission energy allocation [11]. We chose to demonstrate the power of multimode image coding in conjunction with the following techniques: (i) channel-optimized quantizers (COQ) [4] and (ii) rate compatible punctured convolutional codes (RCPC) [5]. We note in passing that further improvements were obtained by incorporating transmission energy allocation methods, but space does not permit us to adequately develop that method here.

4.1 MIC1: with Channel Optimized Quantizers

The objective of channel optimized quantizer design is to optimize the encoder/decoder pair for the given source, channel condition, and bit allocation (rate). This is achieved by modifying the generalized Lloyd algorithm (GLA) to take into account the effect of channel errors (see e.g. [4]).

We design channel matched quantizers, along with the corresponding bit allocation strategy, for the DCT coefficients in the M classes. Let ϵ be the bit error rate on the given binary symmetric channel. We impose the requirement that no coefficient be encoded using more than r_{max} bits.

The probability densities of the DCT coefficients can be reasonably approximated by gaussian distributions [9]. Hence, 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 $\{Q_1^0, Q_2^0, ..., Q_{r_{max}}^0\}$ 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 by scaling Q_r^0 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. The corresponding steps in the iterative design algorithm are:

1. Given the current partition, design the quantizers and bit allocation for each class:

   For each coefficient $i = 1, 2, ..., k$, and for each class $j = 1, 2, ..., M$:

   (a) Compute σ_{ij}^2, the variance of the ith coefficient of blocks belonging to class j.

   (b) Determine bit allocation as $r_{ij} = \arg\min\{d_{ij}(r) + \lambda r\}$, where $d_{ij}(r) = d(r)\sigma_{ij}^2$.

   (c) The quantizer Q_{ij} is obtained by scaling the normal quantizer by σ_{ij}.

2. Redesign the prefix code for class indices c_j based on the empirical rate of occurrence of each class. (including redundancy for heavy protection).

3. Given the new quantizers and prefix code, repartition the training data into classes. For each block the classification decision minimizes
   $\sum_{i=1}^{k}[d_{ij}(r_{ij}) + \lambda r_{ij}] + \lambda c_j$.

4.2 MIC2: Multimode Image Coding with RCPC

Rate Compatible Punctured Convolutional Codes (RCPC) [5] provide unequal error protection while sharing the same encoder/decoder structure enabling them to share the same hardware in a practical realization. We make use of RCPC to tailor the protection offered to each individual bit of the quantized DCT coefficients..

We optimize our source quantizers for the noiseless channel. We design a set of scalar quantizers of rates $r = 1, 2, ..., r_{max}$ bits, where each quantizer is optimized for a unit variance Gaussian source and the noiseless channel. Let us denote these quantizers by $\{Q_1^0, Q_2^0, ..., Q_{r_{max}}^0\}$ and the corresponding distortion they produce by $\{d_s(r)\}$ where the subscript s indicates that the distortion is due to source coding alone. We repeat the design procedure of section 4.1 producing a set of class indices, bit allocation and quantizers for the coefficients.

Next, we evaluate the channel error sensitivity of each bit of the quantized DCT coefficients to apply judicious unequal error protection. We first consider a n-bit quantizer codebook $C = \{y_0, y_1, ..., y_{2^n-1}\}$ and denote the codebook index: $I = (i_1 i_2 ... i_n)$. The sensitivity S_j of the j the bit is defined as the expected amount of distortion caused by a bit error at this location:

$$S_j = E||y_{i_1,i_2,..,i_j,..,i_n} - y_{i_1,i_2,..,i_j^*,..,i_n}||^2 \tag{6}$$

where, $i_j^* = 1 - i_j$ is the complement of i_j.

Let the quantizer encode a DCT coefficient x to y_i and let the decoded codeword be y_j. The average distortion is given by

$$D = E||x - y_j||^2 = E||x - y_i||^2 + E||y_i - y_j||^2 \tag{7}$$

since $\{y_i\}$ are at the centroids of their encoding regions. The first term is the source coding distortion D_s (which is fixed) while the second term can be interpreted as the channel distortion D_c. If we assume that the error rate is small enough that probability of more than a single bit error can be neglected, then

$$D_c = \sum_{i=1}^{n} \epsilon_i S_i \tag{8}$$

where ϵ_i is the error probability and S_i is the sensitivity of each bit. Note that the contribution of each bit to the Lagrangian cost can be minimized independently under this assumption.

We consider a particular RCPC family [7]. Let each code be characterized by the triplet (r_c, r_s, ϵ') where r_c is the channel rate, r_s is the source rate and ϵ' is the decoded error rate produced by the code for the given channel error rate ϵ. Let the sensitivity of a particular bit be denoted by S. Assign the error correction code which minimizes the Lagrangian contribution,

$$\epsilon' S + \lambda * \frac{r_c}{r_s} \tag{9}$$

This is repeated for each bit of the quantized DCT coefficients.

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 MIC1 and MIC2 , with the number of modes $N = 1, 4$ and 16 in both the cases. As discussed, MIC1 uses channel optimized quantizers, while MIC2 uses rate compatible punctured convolutional codes for error protection. Note that the single mode ($N = 1$) cases in MIC1 and MIC2 corresponds to the fixed

R		0.4	0.5	0.6	0.75	1.0
Method	N					
$MIC1$	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
$MIC2$	1	23.61	23.39	23.69	23.85	24.44
	4	26.47	26.72	27.15	28.16	28.79
	16	27.10	27.50	27.74	29.21	29.48

Table 1: 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],[3].

R		0.4	0.5	0.6	0.75	1.0
Method	N					
$MIC1$	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
$MIC2$	1	21.90	21.75	22.07	22.16	22.64
	4	24.84	24.98	25.35	26.05	26.48
	16	25.86	26.10	26.80	27.21	27.50

Table 2: 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],[3].

rate coder of [1] and [2] respectively. 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 [8]. 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 listed rates include the rate required for transmitting the protected mode information. 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.

5 Conclusion

We have proposed a new multimode coding framework to design robust compression schemes for transmission through noisy channels. We developed a design algorithm which jointly optimizes the compression performance and error resilience and applied the algorithm to develop ADCT-based multimode image coding schemes (MIC).

The framework allows the incorporation of various channel protection schemes and we chose to demonstrate its power by incorporating COQ [4] and RCPC [5], yielding the specific coding methods MIC1 and MIC2, respectively. Simulations demonstrate that the multimoder coders can provide significant improvements in performance of up to 6dB over the known fixed-length predecessors [1], [2].

Current work involves extending this framework to handle video coding for noisy channels, and results will be presented in a future publication.

References

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

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

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

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

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

[6] W. H. Chen and C. H. Smith, "Adaptive coding of monochrome and color images," *IEEE Trans. on Com.*, vol. 25, pp. 1285-1292, Nov. 1977.

[7] L. H. C. Lee, "New rate-compatible punctured convolutional codes for Viterbi decoding," *IEEE Trans. on Com.*, vol.42, pp. 3073-9, Dec. 1994.

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

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

[10] M. Effros and P. A. Chou, "Weighted universal bit allocation: optimal multiple quantization matrix coding," *Proc. ICASSP*, pp. 2343-46, 1995.

[11] S. Gadkari and K. Rose, "Robust vector quantisation by transmission energy allocation," *Electronics Letter*, vol. 32, no. 16, pp. 1451-53, Aug. 1996.

[12] K. Zeger and A. Gersho, "Pseudo-Gray coding," *IEEE Trans. on Com.* vol. 38, pp. 2147-58, Dec. 1990.