

# Region-of-Interest Data Compression with Prioritized Buffer Management (II)

Sam Dolinar,  
Aaron Kiely, Matt Klimesh,  
Shervin Shambayati  
Jet Propulsion Laboratory  
California Institute of Technology  
Pasadena, California  
{sam, aaron, klimesh,  
shervin}@shannon.jpl.nasa.gov

Antonio Ortega, Sang-Yong Lee,  
Phoom Sagetong, Hua Xie  
Department of Electrical Engineering  
Signal and Image Processing Institute  
Integrated Media Systems Center  
University of Southern California  
Los Angeles, California  
{ortega, huaxie}@sipi,  
sagetong@scf-fs,  
sangyong@biron}.usc.edu

Roberto Manduchi  
Computer Engineering Department  
University of California, Santa Cruz  
Santa Cruz, California  
manduchi@soe.ucsc.edu

*Abstract* — We describe the second year of work on an integrated system for intelligent compression and transmission of copious data acquired by spaceborne instruments. At its core, our system contains a wavelet-based progressive image compression algorithm, ICER. Our modified version, ROI-ICER, accepts input priorities measuring the relative importance of various “regions of interest” in the source data, and tags its output packets to reflect both the regional priorities and the wavelet bit layer priorities.

The output of the data compression module is supervised by an intelligent buffer manager that receives prioritized packets from many different source images and tries to select packets for transmission that will maximize the total science value of the data received on the ground. Our baseline buffer manager uses a simple form of double-valued prioritization: admissions and discards are determined by priorities established by ROI-ICER, while transmissions are first-in, first out (FIFO) among packets that survive the admission/discard process during their time of residency in the buffer.

Extensions of the baseline classification and prioritization algorithms now cover more realistic earth science scenarios, including applications with multispectral data. Extensions to ROI-ICER incorporate a new data model and compression engine. Improvements to the buffer manager handle dynamically changing priorities. A buffer state parameter can be fed back to save ROI-ICER from performing unnecessary computations.

Various theoretical advances are being evaluated for inclusion in the mainline software: algorithms to optimize quantizers for feature compression and classification for prioritization feedback, based on a criterion trading off rate, distortion and complexity; improvements to the Mallat distortion model that yield better analytical model-based bit allocations for optimizing region-of-interest coding; and a new buffer control criterion that can approximately match both the minimum worst-case distortion achieved by a minimax criterion and the minimum average squared distortion achieved by a minimum mean squared error criterion.

---

This work was funded by the ESTO Technology Program and performed at the Jet Propulsion Laboratory, California Institute of Technology, at the Signal and Image Processing Institute, Integrated Media Systems Center, University of Southern California, and at the Computer Engineering Department, University of California, Santa Cruz.

## I. THE BASELINE SYSTEM

We are in the second year of a three-year project to develop integrated data compression and buffer management algorithms to maximize the science value of data returned from spacecraft instruments [1]. Our approach is to adapt existing progressive compression algorithms to make use of identified “regions of interest” (ROIs) in the data, and to develop buffer strategies for prioritizing, storing, and delivering the most valuable compressed segments, and later reconstituting the original data. Our system incorporates ROI considerations across many images or different data types. The algorithms are subject to practical limits on the onboard computer’s speed, memory, and storage.

At its core, our data compression system contains a wavelet-based progressive image compression algorithm, ICER [2], that is being used on the Mars Exploration Rover (MER) mission. The ICER algorithm applies a wavelet decomposition and prioritizes the compressed bit layers from the wavelet subbands so as to progressively transmit the layer that gives the largest estimated improvement in image quality per transmitted bit. Our modified version, ROI-ICER, accepts additional input priorities in the form of a data prioritization map that gives the relative importance of different regions of interest in the source data. Then ROI-ICER produces output packets of compressed data along with priority labels that reflect both the input regional priorities and the wavelet bit layer priorities.

The output of the data compression module is supervised by an intelligent buffer manager that receives prioritized packets from many different source images and tries to select packets for transmission that will maximize the total science value received on the ground. Just as importantly, it attempts to discard only the least valuable packets when the buffer overflows (which is inevitable if the average data transmission rate is lower than the average data collection rate). Our buffer manager uses a simple form of double-valued prioritization: admissions and discards are determined by priorities established by ROI-ICER, while transmissions are first-in, first out (FIFO) among packets that survive the admission/discard process during their time of residency in the buffer. The FIFO protocol for transmissions keeps intact the chains of compressed data packets that are later used to progressively reconstruct each image or image segment, yet the prioritized decisions on admissions and discards ensure that the scarce downlink resource is not clogged by less valuable data. Using a FIFO transmission priority eliminates the need to unshuffle the packets received on the ground, because successive (truncated) packet chains can be used to reconstruct the source images in the same order in which they were acquired (but to different levels of distortion depending on how many packets from each chain survived

the prioritized admission/discard process).

## II. IMPROVEMENTS TO THE BASELINE SYSTEM

In this section we describe several recent improvements to our baseline system. Extensions of the baseline classification and prioritization algorithms now cover more realistic earth science scenarios, including applications with multispectral data. Extensions to ROI-ICER incorporate a new data model and compression engine. Improvements to the buffer manager apply the double-valued priority queue model to handle dynamically changing priorities, and permit feedback of a buffer state parameter to save ROI-ICER from performing unnecessary computations.

### A. DATA CLASSIFICATION AND PRIORITIZATION

Our baseline (color-based) image classification algorithm is being extended to cover more interesting or sophisticated earth science classification scenarios. Prioritization is particularly effective for missions that are well-targeted in terms of data usage and science goals. One such application is crop monitoring using multispectral aerial imagery [3, 4]. This is an area of intense current research, given its potential for agricultural applications. We are looking at an ongoing NASA-sponsored project [5] for coffee harvesting optimization using imagery from airborne cameras on a solar-powered unmanned aerial vehicle (UAV). The UAV has two multispectral (3-band), high-resolution ( $4072 \times 4072$  pixels) cameras. A sample image is shown in Fig. 1. This project<sup>1</sup> represents a typical scenario



Figure 1: Example of a coffee field image from the UAV.

where priorities can be assigned to image segments in a rather unambiguous manner.

For this application a two-stage classification system is effective. A first-stage classifier based on color, texture, and shape can identify the large-scale areas of interest (i.e., the coffee orchard areas) and distinguish them from uninteresting areas such as roads and tracts

<sup>1</sup>More information about this project, led by Stan Herwitz at Clark University, can be found at <http://www.clarku.edu/research/access/geography/herwitz/herwitzD.shtml>. We are grateful to Lee Johnson of California State University, Monterey Bay, and NASA Ames Research Center, for providing us with information about the project and with sample aerial pictures (including the image in Fig. 1) taken during preliminary tests in Kauai.

of houses. Then a second-stage classifier with additional fine-detail recognition based on multispectral and geometric analysis can identify individual rows of plants by distinguishing pixels belonging to the sunlit canopy from pixels corresponding to soil or shaded vegetation. The results of such a two-stage classification scheme applied to the image in Fig. 1 are shown in Fig. 2. The information from the

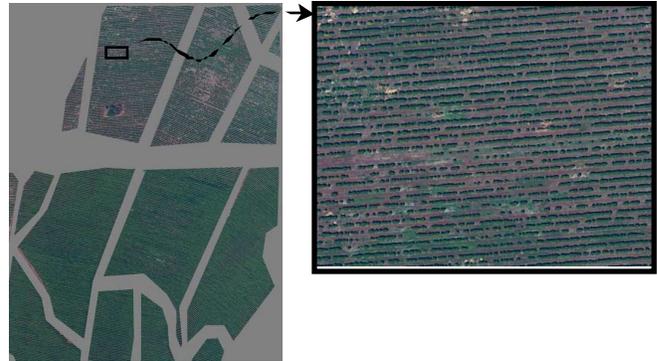


Figure 2: Two-stage prioritization scheme for coffee field image: a first-stage classifier identifies orchard areas, then a second-stage classifier identifies the sunlit canopy within the orchards.

sunlit canopy provides the best data for characterizing the state of the coffee plants, and should be given the highest priority. The information from the soil and the shaded vegetation is less useful, and is assigned an intermediate priority.

The image prioritization algorithms for this scenario exploit a number of different visual cues. We are using both pixel-wise photometric information from the 3-band cameras and spatial pattern analysis. Vegetation indices (computed by suitable ratios of the color value in different channels) are useful to determine whether a pixel belongs to a vegetated surface, although they are not very reliable in the case of shadows or inter-reflections. Texture feature segmentation and line or edge detection may be very powerful for determining the boundaries of orchards, which typically appear as regular spatial patterns. To combine the two types of visual features (photometric and textural), we are developing robust statistical algorithms based on the Bayesian fusion of classifiers [6]. A characteristic of this approach is that independent classifiers can be developed for different features (allowing one to exploit physical or statistical models that are peculiar to each feature), and then combined together according to a simple yet statistically sound rule.

### B. DATA COMPRESSION

Sophisticated strategies are being developed to handle multispectral images with identified regions of interest. As described in [1], the baseline ROI-ICER exploits correlations in 3-band color RGB images by transforming them to the YCrCb domain. The generalization of this procedure for multispectral data is to apply a decorrelating transform in the spectral dimension, then process the transformed components independently. With multispectral data it is also important to generalize the notion of regions of interest to cover not just spatial regions but also regions of interest in the frequency dimension, assigning higher priorities to some spectral bands

Many recent improvements to the ICER compression software are being implemented in the new version of ROI-ICER. ICER has been

re-engineered to provide improved compression effectiveness in a number of ways [2]. It incorporates a sophisticated data model and compression engine to more effectively compress the wavelet transformed data. It also includes a reorganized output data structure that simplifies the implementation of *error containment* strategies to limit the effects of packet losses.

ICER employs a technique known as *context modeling* in its encoding of the bit planes of wavelet-transformed data. Before encoding a bit in the transformed image, the bit is classified into one of several contexts based on the values of previously encoded bits. These bits are from the pixel being encoded and nearby pixels, and include the bits previously encoded from the current bit plane as well as bits from previous (more significant) bit planes. The probability that the bit to be encoded is a “0” is estimated based on the encoder’s previous experience with bits classified into the same context. The bit is then encoded based on this probability estimate, as described below. Since the probability estimate relies only on previously encoded information, the decoder can duplicate this calculation and produce the same probability estimate, which is essential for proper decoding. ICER uses a reasonably low-complexity scheme for classification into contexts: it needs only two small lookup tables, is based on simple properties of the pixel being encoded as well as the 8 nearest-neighbor pixels, and yields one of 18 contexts. This scheme is similar to the context model used by JPEG 2000 [7].

Compression of the bits is accomplished with an adaptable entropy coder. The entropy coder takes the sequence of bits to be encoded, along with probability estimates obtained from the context modeler, and produces a compressed bit stream from which the original sequence of bits can be reconstructed. ICER uses a technique called interleaved entropy coding [8, 9, 10] that has the same functionality as arithmetic coding, which is the current state-of-the-art in adaptable entropy coding (used, for example, in JPEG 2000 [11]). Both methods achieve excellent performance (within 1% of the theoretical limit in typical applications), but the interleaved entropy coder can be implemented with particularly low complexity, and so is well-suited for space applications where speed is of critical importance.

Error containment can provide some protection against corruption from packet losses that arise on the communications channel. Without error containment, a single packet loss due to channel errors can corrupt large segments of compressed data. To achieve error containment, ICER produces the compressed bitstream in separate pieces or segments that can be decoded independently. These segments represent rectangular regions of the original image, but are defined in the transform domain<sup>2</sup>. ICER provides flexibility in choosing an appropriate number of error containment segments, since this choice involves trading off compression effectiveness against packet loss protection, thereby accommodating different packet loss rates. Dividing an image into a large number of segments can confine the effects of a packet loss to a small area of the image, but it is generally harder to compress smaller image segments effectively. However, increas-

<sup>2</sup>The alternative method, to apply the wavelet transform separately to each segment of a partitioned image, is less desirable because the boundaries between segments would be noticeable in the reconstructed image after lossy compression even when no packet losses occur. By segmenting the image in the transform domain, we can virtually guarantee that such artifacts will not occur. There are also secondary benefits: we achieve better decorrelation by applying the wavelet transform to the entire image at once, and it is easier to maintain a similar image quality in the different segments. A minor side effect is that the effect of data loss in one segment can appear to “bleed” slightly into adjacent segments.

ing the number of segments can improve compression effectiveness when disparate regions of the image end up in different segments.

The current version of ICER does not attempt to draw the partition boundaries so that they contain disparate regions. Small improvements in the compression efficiencies of individual segments may be offset by the increased overhead needed to describe complex segmentations. With ROI-ICER, however, there is increased motivation to match the partitioning of the image for error containment purposes to an (approximate) partitioning imputed by the classification algorithms, so we will have to study in more detail the tradeoffs that arise from increasing the complexity of segmentation.

### C. BUFFER MANAGEMENT

We ran simulations under different loading conditions to evaluate the performance of the baseline buffer manager. The simulations modeled Poisson or constant-rate arrivals of compressed packet chains of varying lengths from a set of test images, together with a constant transmission rate. Fig. 3 shows an example of a histogram of priorities of packets transmitted, blocked, or discarded from the buffer when the buffer was presented with a load factor of 6 Erlangs (an average of 6 packets arriving for each packet transmitted). We found that, for all loading conditions simulated, the buffer man-

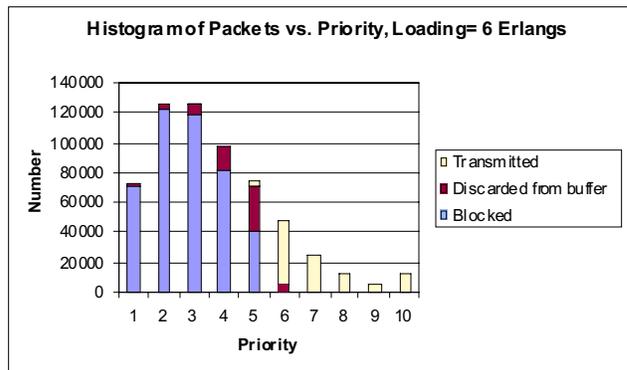


Figure 3: Example of the buffer manager’s performance.

ager almost always transmits the most important data permitted by the loading. In the example of Fig. 3, the cutoff priority for transmission, determined by the load factor, is approximately at priority value 6. In this case, only a small number of packets with priority 6 are discarded, and only a small number of packets with priority 5 are transmitted. Our double-valued priority queue is very efficient at sifting through all the packets and making sure that the downlink is filled with only the highest-value packets that fit. Furthermore, we see from Fig. 3 that most of the untransmitted packets are “blocked” before they can ever be admitted to the buffer. The relatively few packets discarded after admission are generally closer in priority to those transmitted. This indicates that our baseline buffer manager’s computational resources are devoted mostly to the borderline decisions, while the easy decisions are dispatched quickly.

This efficiency can be carried one step further if the buffer manager feeds back information to ROI-ICER on its current cutoff priority value for admissions. There is little reason for ROI-ICER to devote its computational resources to producing low-priority packets that will almost certainly be discarded. ICER itself has a new built-in capability to stop computing compressed data packets when an “image quality” requirement is satisfied. This feature will allow

ROI-ICER to monitor the state of the buffer and stop producing the lowest priority packets.

Another measure of the buffer manager’s performance is its transmitted packet delay distribution, as illustrated in Fig. 4. Delay is the

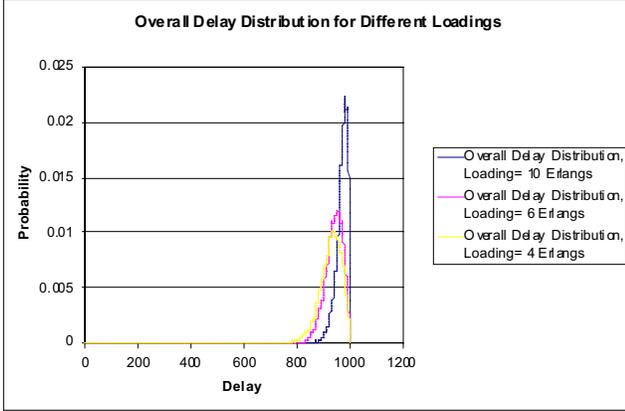


Figure 4: The buffer manager’s packet delay distribution.

time between arrival and transmission of transmitted packets. Under a FIFO transmission protocol, the maximum possible delay for every transmitted packet equals the size of the buffer divided by the (constant) transmission rate (denoted arbitrarily by 1000 in the figure). Under the prioritized discard rule, a given packet’s transit time through the buffer is reduced every time a lower priority packet is ejected from a position further ahead in the queue. However, if a relatively small fraction of packets are discarded after admission (as in Fig. 3), then the delay for most packets will be close to the maximum and nearly constant (as in Fig. 4).

The buffer manager’s primary policing job is to ensure that unworthy packets have little opportunity to sneak into the downlink channel during a period of very few high-priority arrivals. With a plain FIFO transmission protocol, low-priority packets will necessarily be transmitted if a packet’s transit time through the buffer is less than the duration of an extended series of low-priority arrivals. Thus, it is important to make the buffer large enough that a packet’s residency time under the FIFO transmission protocol is long enough to average out the normal fluctuations in the priorities of arriving packets.

A modification of the FIFO transmission protocol can allow the buffer manager to respond to dynamically changing priorities. The buffer manager can accomplish this by setting aside, but not discarding, data segments whose ultimate value is still in limbo at the time they would normally reach the front of the transmission queue. Packets routed to the secondary “set-aside” queue can rest there until more information is available to decide their fate. The motivation for establishing the set-aside queue is to preserve the simplicity of the FIFO queue for the vast majority of packets that require no reevaluation. Effective operation of the set-aside queue requires much more intricate logic than that governing the FIFO queue. Packets with significantly changeable priorities should be flagged so they can be yanked from the FIFO queue or rescued from the trash. Such packets constitute the set-aside queue’s arrivals. Departures from the set-aside queue occur when a packet’s priority is deemed to be final. Such departures can be to the trash, to the downlink, or to the back of the FIFO queue. The last option eliminates any need for the primary

FIFO queue to time-share the downlink with the secondary set-aside queue.

### III. ADDITIONAL THEORETICAL ADVANCES

Various theoretical advances are being evaluated for inclusion in the mainline software: algorithms to optimize quantizers for feature compression and classification for prioritization feedback, based on a criterion trading off rate, distortion and complexity; improvements to the Mallat distortion model that yield better analytical model-based bit allocations for optimizing region-of-interest coding; and a new buffer control criterion that can approximately match both the minimum worst-case distortion achieved by a minimax criterion and the minimum average squared distortion achieved by a minimum mean squared error criterion.

#### A. ENTROPY- AND COMPLEXITY-CONSTRAINED CLASSIFIED QUANTIZATION DESIGN FOR IMAGE CLASSIFICATION<sup>3</sup>

Our previous work [1, 12] concentrated on feature compression and classification for prioritization feedback. We proposed to represent image regions by a feature set and then to compress this feature set and to transmit it through the downlink to the ground. A priority is assigned to the corresponding region by performing a search in the database using the compressed feature. We showed that, rather than simply compressing the features for full classification on the ground, it was more efficient to perform some pre-classification onboard before transmitting the features. The main benefit from this pre-classification was to enable different quantizers to be designed specific to each leaf of the pre-classifier’s decision tree, and thus better matched to the characteristics of the data.

Several criteria should be kept in mind to design the compression and classification system. Efficiency of a compression scheme is measured not only by its rate-classification performance, but also by its conservation of onboard processing power, i.e., its coding complexity. In our new work we show how to optimally split the classification tree between onboard and ground processing, based on a tradeoff of complexity, rate, and distortion constraints.

A *decision tree classifier* is applied to classify the compressed data. We assume that the pre-classifier is a pruned subtree of the original decision tree. The Generalized Breiman, Freidman, Olshen, and Stone (G-BFOS) algorithm is employed to jointly search for the optimal pre-classifier and quantization parameters for each of the classes. The optimization is carried out based on not only the rate budget, but on the coding complexity constraint as well. Previously we illustrated this framework by showing a texture classification example. Although we showed an example where a simple uniform scalar quantization is employed and a  $K$ -means tree is used as the classifier, the idea is easily generalized to any classified quantization system where a decision tree is involved.

The goal is to find the optimal subtree  $S^* \leq \mathcal{T}$  and the set of stepsizes  $\{\Delta_{i,j}^*, j = 1, \dots, N\}$  for each class  $i$ , such that the overall distortion  $D^*$  is minimized subject to the rate budget  $R_b$  and complexity constraint  $C_b$ .

$$D^* = \min_{S^*, \{\Delta_{i,j}^*\}} \sum_{i=1}^{|\tilde{S}|} P_i \times D_i(\Delta_{i,1}, \Delta_{i,2}, \dots, \Delta_{i,N}) \quad (1)$$

such that  $R(S, \{\Delta_{i,j}\}) \leq R_b$  and  $C(S) \leq C_b$

<sup>3</sup>Work published in part in [13].

Let us introduce some notation.  $D()$ ,  $C()$  and  $R()$  are tree functionals defined as:

$$\begin{aligned} D(S, \{\Delta_{i,j}\}) &= \sum_{t \in \tilde{S}} P(t) \times d(t) \\ R(S, \{\Delta_{i,j}\}) &= \sum_{t \in \tilde{S}} P(t) \times r(t) \\ C(S, \{\Delta_{i,j}\}) &= \sum_{t \in \tilde{S}} P(t) \times l(t) + w \times |\tilde{S}| \end{aligned} \quad (2)$$

where  $\tilde{S}$  is the set of leaf nodes of  $S$  and its size equals the number of encoders that need to be stored;  $P(t)$  is the probability that an input vector traverses node  $t$ ;  $l(t)$  is the length of the path from the root to node  $t$ , and reflects the cost of traversing the classification tree  $S$  from the root to node  $t$ ;  $w$  is a positive weighting factor; and  $r(t)$  and  $d(t)$  are the operating entropy rate and distortion, respectively, of the quantized output for the sample space at node  $t$ , with the set of quantization stepsizes  $\{\Delta_{i,j}\}$  applied to quantize the data at node  $t$ :

$$\begin{aligned} r(t_i, \{\Delta_{i,j}\}) &= \sum_{X \in t_i} \sum_{k=1}^N H(\hat{x}_k) \\ d(t_i, \{\Delta_{i,j}\}) &= \sum_{X \in t_i} \sum_{k=1}^N d(x_k, \hat{x}_k) \end{aligned} \quad (3)$$

Instead of solving the constrained problem (2), we use Lagrange multipliers and solve the dual problem:

$$\min_{S^*} [\min_{\{\Delta_{i,j}^*\}} \{D(S, \{\Delta_{i,j}\}) + \lambda \times R(S, \{\Delta_{i,j}\}) + \mu \times C(S)\}] \quad (4)$$

Now we need to find the optimal multipliers  $\lambda$  and  $\mu$  such that the rate and complexity constraints are satisfied with equality.

The BFOS algorithm proposed by Friedman et al. [14] is a Lagrangian flavored method. It minimizes the functional  $J(S) = \delta(S) + \lambda \times l(S)$  over all pruned subtrees  $S \leq \mathcal{T}$ , with  $\delta(S)$  and  $l(S)$  being the average distortion and rate of a tree structured vector quantizer defined on  $\mathcal{T}$  and pruned to  $S$ . Chou et al. [15] extended the BFOS algorithm by generalizing the two components of the cost functional to any tree functionals  $U_1(S)$  and  $U_2(S)$ . It was proved that the generalized BFOS (G-BFOS) algorithm is capable of tracking out the extreme points which lie on the convex hull of the operating  $(U_1, U_2)$  pairs over all possible pruned subtrees  $S \leq \mathcal{T}$ . We define the two tree functionals as follows:

$$\begin{aligned} U_1(S, \{\Delta_{i,j}\}) &= D(S, \{\Delta_{i,j}\}) + \lambda \times R(S, \{\Delta_{i,j}\}) \\ U_2(S, \{\Delta_{i,j}\}) &= C(S, \{\Delta_{i,j}\}) \end{aligned} \quad (5)$$

As the sample space is hierarchically partitioned by the tree  $\mathcal{T}$ , within each node  $t$  data tends to be clustered. This property can be exploited to enable efficient quantization and entropy coding. Due to this coding gain, for a fixed multiplier  $\lambda$ , the tree functional  $U_1$  would be monotonically decreasing as the tree grows. On the other hand, the complexity of the system is monotonically increasing since both the depth of the tree and the number of encoders become larger. This ensures that an optimal point will be found by pruning the tree. We proposed a nested optimization algorithm to jointly search for the optimal subtree  $S^*$  and the set of quantization stepsizes  $\{\Delta_{i,j}\}$  for a given rate budget and complexity constraint.

The basic idea is: Initialize the multiplier  $\lambda$ , then for this fixed multiplier, choose the set of stepsizes which minimize the functional

$u_1(t) \forall t \in \mathcal{T}$ . Then the G-BFOS algorithm is used to prune the tree until the complexity constraint  $C_b$  is satisfied. Then given the resulting operating rate  $R$ , we adjust the multiplier  $\lambda$  using the bisection method [16] and repeat the process until convergence. We perform the optimization on a binary tree  $\mathcal{T}$ .  $\mathcal{T}$  is built in a top-down manner using the  $K$ -means algorithm until only one class is left at each leaf node. The training set  $L = \{X, Y\}$  is a labeled vector sequence obtained by computing the wavelet feature of the Brodatz texture album [12].  $X$  is the feature vector and  $Y$  is the associated texture label. The stepsizes  $\{\Delta_{i,j}\}$  are chosen from the predefined discrete set  $\{32, 16, 8, 4, 2\}$ . We assume the traditional mean squared error distortion, but it can be extended to other distortion measures. Figure 5 shows the obtained Rate-Distortion-Complexity surface. We can see the tradeoff between rate, distortion, and complexity of the system, and we can verify the convexity of the surface, as predicted by Goyal and Vetterli [17]. In Figure 6, we show the rate-distortion performance for this coding system with and without pre-classification. Substantial coding gain is obtained by employing a classified quantizer instead of a single quantizer, or by using a higher-complexity classification tree.

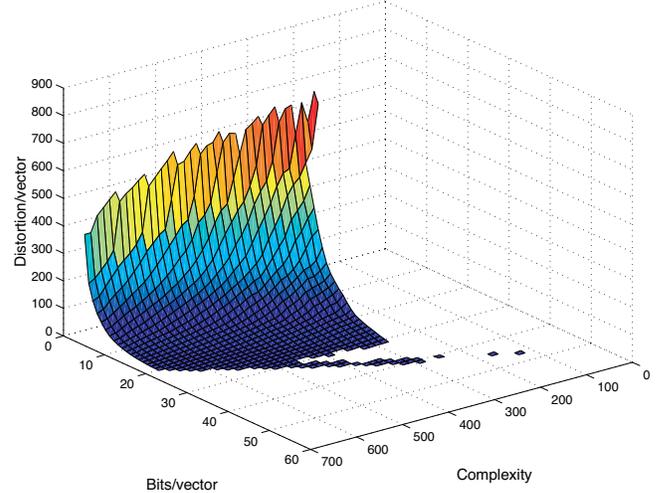


Figure 5: The Rate-Distortion-Complexity surface obtained by the proposed optimization framework.

We also performed a texture classification with the compressed data and the result is shown in Figure 7. A lower classification error rate was achieved by using a classified encoder instead of a single encoder.

#### B. ANALYTICAL MODEL-BASED BIT ALLOCATION FOR OPTIMIZATION OF REGION OF INTEREST CODING <sup>4</sup>

In our previous work [1, 18] we showed how one could apply the Mallat model [20] to solve the problem of optimal bit allocation in a region-of-interest setting [18]. We have now modified this model to better take into account the special characteristics of the problem, in particular ensuring that the model fits the data well at both high and low rates.

<sup>4</sup>Work published in part in [18, 19].

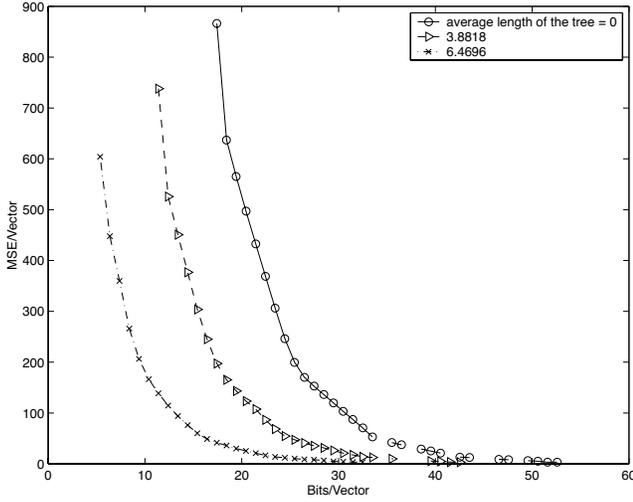


Figure 6: Comparison of rate-distortion performance between systems with pre-classification (using tree lengths 3.8818 and 6.4696) and without pre-classification (tree length 0).

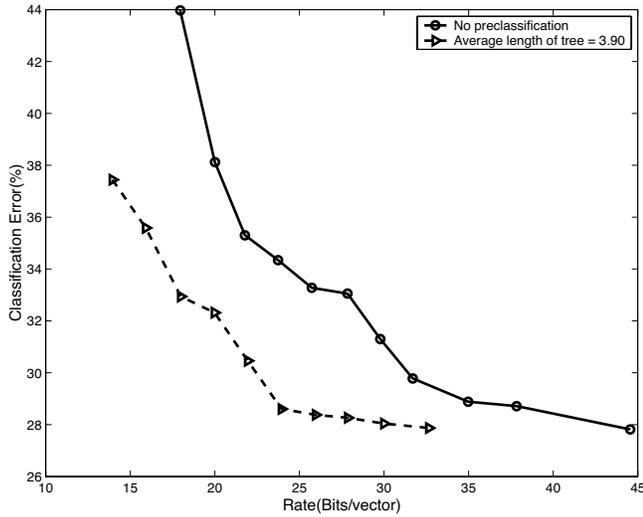


Figure 7: Classification performance for systems with and without pre-classification.

Our objective is to allocate bits to different regions in an image coded with a progressive wavelet coder such as ICER [2], SPIHT (Set Partitioning in Hierarchical Trees) [21] or JPEG 2000 [11], in order to achieve region-of-interest (ROI) coding objectives. Here we introduce a novel Rate-Distortion (R-D) model which is an extension of Mallat's model and especially designed to capture R-D behavior when different parts of an image are refined at different speeds. Our model takes into account that the rates used are not necessarily the same throughout the image. Because of the different rates, certain modeling approximations (e.g., those for coarse quantization) can not be used uniformly throughout the image.

First, we briefly discuss the R-D model for SPIHT originally proposed by Mallat and Falzon [20]. Given a total bit rate budget for an image  $R_b$ , the average quantization error  $D(M_b)$  is the summation of the quantization error due to quantizing the significant coefficients  $E^{sig}$  and that due to setting to zero the insignificant coefficients  $E^{insig}$ , divided by the total number of wavelet-transformed coefficients,  $N$ :

$$\begin{aligned} D(M_b) &= \frac{E^{sig} + E^{insig}}{N} \\ &= \frac{M_b \frac{\Delta^2}{12} + \sum_{i=M_b+1}^N |x(i)|^2}{N}, \end{aligned} \quad (6)$$

where  $M_b$  is the number of significant coefficients with amplitude larger than  $\Delta$ , the size of final quantization bin, i.e., the final threshold. Also,  $M_b$  is directly related to the average bit rate  $R_b$  via  $M_b = \frac{NR_b}{6.6}$ . All  $N$  wavelet coefficients are sorted in monotonically descending order of magnitude, to obtain a list  $\{x(i), i = 1, \dots, N\}$ .  $E^{insig}$  is the energy of the  $N - M_b$  smallest amplitude coefficients since it is the error when all insignificant coefficients are quantized to zero, i.e.,  $E^{insig} = \sum_{i=M_b+1}^N |x(i)|^2$ . According to the sorted sequence, it is clear that  $\Delta = |x(M_b)|$ . The average quantization error per significant coefficient is calculated based on the hypothesis that the probability density function (pdf) of the significant coefficients is flat within each quantization interval and thus the well-known approximation of a uniform distribution can be used,  $\frac{E^{sig}}{M_b} = \frac{\Delta^2}{12}$ . As shown in Figure 8, this explains why Mallat's model works well at low bit rates. This is because the histogram outside the central bin,  $\Delta_L^o$ , is sufficiently flat, with coarse quantization, leading to an accurate approximation for  $E^{sig}$ . However, at high bit rates, the histogram outside  $\Delta_H^o$  is not very flat, so that approximating it by a uniform distribution within the interval may not be accurate. Even though at high rates the quantization bins are small, they are not sufficiently small to make the uniform approximation sufficiently accurate.

In our proposed modified model, we start by assuming that the pdf of the wavelet coefficients can be modeled as a Laplacian distribution. We then obtain a new representation of the error for the significant coefficients which is given by:

$$E^{sig} = \frac{M_b e^{-K} \left(1 - e^{-\frac{KS}{2}}\right) \left(K^2 - 4K + 8 - e^{-K} (K^2 + 4K + 8)\right)}{4\lambda^2 (1 - e^{-K})}, \quad (7)$$

where  $K = \lambda\Delta$  and  $S$  is the number of quantization intervals besides the central bin. Note that  $E^{insig}$  is unchanged since there is no approximation in representing the error due to setting to zero the insignificant coefficients. More detailed descriptions can be found in [19].

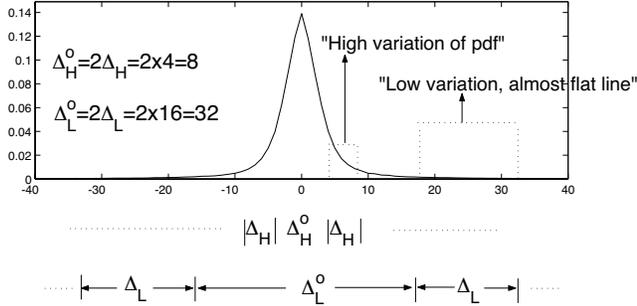


Figure 8: Normalized histogram of the wavelet coefficients of the gray-level Lena image.  $\Delta_L = 16$  and  $\Delta_H = 4$  are the sizes of the final quantization bin and,  $\Delta_L^o = 32$  and  $\Delta_H^o = 8$  are the sizes of the zero bin, at low and high bit rates (0.3 and 1.1 bps), respectively.

In our experiments, we applied the distortion ratio criterion to three standard gray-level images of size  $512 \times 512$  pixels (Lena, Boat and Lake). We validated our results by using different types of ROIs in different positions in each image: (i) a rectangular ROI of size  $200 \times 200$  in the middle of the image, (ii) a cross-shaped ROI in the middle of the image, and (iii) an L-shaped ROI in the upper-left corner of the image. We determine the priority scaling factor ( $psf$ ) obtained from our proposed model, denoted  $psf_{pro}$ , the one obtained using Mallat's model, denoted  $psf_{mal}$ , and the one obtained through an exhaustive search (i.e., selecting the best value among  $psf$  in the set  $\{1, 1.1, 1.2, \dots, 400\}$ ), which we denote  $psf_{emp}$ . Each of the  $psf$  values is chosen so as to target a desired distortion ratio between the ROI and the rest of the image, while minimizing the overall distortion, for a given total rate budget. We computed the variation between  $psf_{pro}$  and  $psf_{emp}$  and the variation between  $psf_{mal}$  and  $psf_{emp}$ . For each shape, we averaged the variations of the means (mean) and standard deviations (std) over 3 images when the desired ratio varied from 1 to 10. We computed 3 statistical sets of data which are: (i)  $psf$  values, (ii) ROI distortions ( $dist_{roi}$ ), and (iii) background distortions ( $dist_{nroi}$ ). These are shown in Tables 1(a) and 1(b), based on our proposed model and on Mallat's model, respectively. Our results show clearly that our proposed model provides more accurate estimates of  $psf$  values and the distortions than those obtained by using Mallat's model.

### C. OPTIMAL RATE CONTROL FOR IMAGE TRANSMISSION OVER CONSTANT BIT RATE CHANNELS BASED ON A HYBRID MMAX/MMSE CRITERION<sup>5</sup>

Our work on rate control of prioritized data under buffering constraints first focused on the use of on-line algorithms for the minimum maximum (MMAX) distortion criterion [1, 22]. For this criterion we showed that an on-line method with buffer sorting is as good as the best off-line approach. We have now extended our work to introduce a new distortion measure criterion, seeking to guarantee good worst-case quality, while also aiming at good average quality. We focus on finding optimal off-line rate control for constant bit rate (CBR) transmission, where the size of the encoder buffer and the channel rate are used as constraints. After the constraints have

<sup>5</sup>Work published in part in [22].

Table 1: Experimental results obtained by using (a) the proposed model and (b) Mallat's model.

Shape	$psf$		$dist_{roi}$		$dist_{nroi}$	
	mean	std	mean	std	mean	std
rectangular	2.23%	1.70%	1.94%	1.64%	1.01%	0.91%
cross	2.83%	1.93%	1.95%	1.90%	1.70%	1.30%
L	2.30%	1.74%	1.74%	1.26%	1.30%	1.12%

(a)

Shape	$psf$		$dist_{roi}$		$dist_{nroi}$	
	mean	std	mean	std	mean	std
rectangular	6.75%	3.12%	2.99%	2.81%	3.45%	2.45%
cross	5.87%	3.01%	3.79%	2.47%	3.39%	2.16%
L	6.14%	3.50%	4.22%	3.75%	2.90%	1.81%

(b)

been determined, a target quality measure should be chosen. Most previous work for image and video coding has been based on minimization of average distortion (MMSE). A main drawback of the MMSE criterion is that the quality difference between images can be large and some images may be coded at relatively low quality even though the average quality is high. The MMAX criterion has been proposed to prevent this heavy fluctuation of source quality. Using this criterion, coding units having a significantly lower than average quality can be avoided. But when multiple constraints are present, as in our case, the MMAX criterion by itself may be inefficient. This is because the MMAX optimization is terminated as soon as it cannot decrease the maximum overall distortion. This means that more data could be sent in these periods and the overall quality can be increased. As an approach to increase overall quality after finding a MMAX solution, we propose to use a MMSE criterion for the remaining bit-budget. We denote this criterion MMAX+, because it augments the MMAX criterion with additional targets.

Table 2 shows the experimental results for each criterion. These results were obtained using a video sequence to provide the sequence of images; we expect comparable results to be achievable when sequences of independent images are used instead. The results show that the proposed MMAX+ criterion gives higher average PSNR than the MMAX criterion while it gives same minimum PSNR.

Table 2: Performance (PSNR) comparison of MMAX, MMAX+ and MMSE optimal solutions. Used channel rate is 10 Mbps, the image interval is 0.5 second and the size of an encoder buffer is 20 Mbits. Initial and final buffer states are at mid-buffer.

Method	Avg.	Std. Dev.	Min.	Max.
MMAX	38.21	0.137	38.15	39.44
MMAX+	38.60	0.847	38.15	42.57
MMSE	38.72	1.424	35.26	42.75

## REFERENCES

- [1] S. Dolinar, A. Kiely, M. Klimesh, R. Manduchi, A. Ortega, S. Lee, P. Sagetong, H. Xie, G. Chinn, J. Harel, S. Shambayati, and M. Vida, "Region-of-Interest Data Compression with Prioritized Buffer Management," 2001 Earth Science Technology Conference, College Park, MD, Aug. 28–30, 2001.
- [2] A. Kiely, M. Klimesh, and J. Maki, "ICER on Mars: Wavelet-Based Image Compression for the Mars Exploration Rovers," *IND Technology and Science News*, Issue 15, to appear, 2002.
- [3] M. S. Moran, Y. Inoue, and E. M. Barnes, "Opportunities and Limitations for Image-Based Remote Sensing in Precision Crop Management," *Remote Sensing of the Environment*, 61:319–346, 1997.
- [4] L. F. Johnson, D. Bosch, D. Williams, and B. Lobitz, "Remote Sensing of Vineyard Management Zones: Implications for Wine Quality," *Applied Engineering in Agriculture*, 17:557–560, 2001.
- [5] S. R. Herwitz, L. F. Johnson, J. C. Arvesen, R. Higgins, J. G. Leung, S. E. Dunagan, "Precision Agriculture as a Commercial Application for Solar-Powered Unmanned Aerial Vehicles," 1st American Institute of Aeronautics and Astronautics UAV Conference, 2002.
- [6] R. Manduchi, "Bayesian Fusion of Visual Classifiers," submitted to the European Conference on Computer Vision, 2002.
- [7] D. Taubman, "High Performance Scalable Image Compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1158–1170, July, 2000.
- [8] P. G. Howard, "Interleaving Entropy Codes," *Proc. Compression and Complexity of Sequences 1997*, Salerno, Italy, pp. 45–55, 1998.
- [9] A. Kiely and M. Klimesh, "A New Entropy Coding Technique for Data Compression," *IPN Progress Report 42-146*, pp. 1–48, April–June, 2001.
- [10] A. Kiely and M. Klimesh, "Memory-Efficient Recursive Interleaved Entropy Coding," *IPN Progress Report 42-146*, pp. 1–14, April–June, 2001.
- [11] M. D. Adams, *The JPEG-2000 Still Image Compression Standard*, ISO/IEC JTC 1/SC 29/WG 1 N 2412, Sept. 2001.
- [12] H. Xie and A. Ortega, "Feature representation and compression for content-based image retrieval," *VCIP*, vol. 4310, pp. 111–122, Jan., 2001.
- [13] H. Xie and A. Ortega, "Entropy- and complexity-constrained quantizer design for distributed image classification," submitted to *MMSP'02*, Dec., 2002.
- [14] L. Breiman, J. H. Freidman, R. A. Olsehn, and C. J. Stone, *Classification and Regression Trees*, Wadsworth, 1984.
- [15] P. A. Chou, T. Lookabough, and R. M. Gray, "Optimal pruning with applications to tree-structured source coding and modeling," *IEEE Trans. on Info. Theory*, vol. IT-35, pp. 299–315, March. 1989.
- [16] K. Ramchandran and M. Vetterli, "Best Wavelet Packet Bases in a Rate-Distortion Sense," *IEEE Trans. on Image processing*, vol. 2, pp. 160–175, April, 1993.
- [17] V. Goyal and M. Vetterli, "Computation distortion characteristics of block transform coding," *Proc. of ICASSP*, Munich, Germany, April, 1997.
- [18] P. Sagetong and A. Ortega, "Analytical Model-Based Bit Allocation for Wavelet Coding with Applications to Multiple Description Coding and Region of Interest Coding," *IEEE 2001 International Conference on Multimedia and Expo (ICME)*, Tokyo, Japan, Aug., 2001.
- [19] P. Sagetong and A. Ortega, "Rate-distortion model and analytical bit allocation for wavelet-based region of interest coding," *IEEE 2002 International Conference on Image Processing (ICIP)*, Rochester, New York, Aug., 2002.
- [20] S. Mallat and F. Falzon, "Analysis of low bit rate image transform coding," *IEEE Trans. on Signal Processing*, vol. 46, no. 4, pp. 1027–1042, April, 1998.
- [21] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 6, no. 4, pp. 243–250, June 1996.
- [22] S.-Y. Lee and A. Ortega, "Optimal rate control for video coding based on a hybrid MMAX/MMSE criterion," submitted to *Proc. Int. Conf. on Image Proc.*, 2002.