

# A TREE STRUCTURED BAYESIAN SCALAR QUANTIZER FOR WAVELET BASED IMAGE COMPRESSION

*Birsen Yazici, Mary L. Comer, R. L. Kashyap, and Edward J. Delp*

School of Electrical Engineering  
Purdue University  
West Lafayette, Indiana 47907-1285

## ABSTRACT

Multiresolution image decompositions (e. g., wavelets), in conjunction with a variety of quantization schemes, have been shown to be very effective for image compression. Recently, several promising tree-structured quantization schemes that exploit the correlation across scales have been proposed. In this paper, we present an image compression algorithm based on a multiresolution Markov random field model used to model the correlations of wavelet coefficients across the scales. We also present experimental results obtained using the algorithm.

## 1. INTRODUCTION

The development of wavelet representations of images has led to efficient image processing and computer vision algorithms. The discrete wavelet transform of an image provides a set of wavelet coefficients which represent the image at multiple scales [1, 2]. Fewer coefficients are required to represent the image at coarse scales than at finer scales. Thus, if the wavelet representation is viewed as a multi-level data structure, with the coarsest scale coefficients at the top level and the finest scale coefficients at the bottom level, then the structure can be viewed as a pyramid. Each node in the pyramid represents one wavelet coefficient.

Wavelet representations have formed a basis for several image compression algorithms [3, 4, 5, 6]. One of the key issues in the development of any wavelet based image compression algorithm is the quantization strategy. Typically, coefficients at different levels of the pyramid are assumed independent and quantized separately [3, 7]. However, this approach does not exploit correlations between coefficients at different levels of

the pyramid. Recently, a number of promising quantizers, known as zerotree quantizers, exploiting correlations between different levels of the pyramid, have been proposed [4, 5, 8]. The key idea in [5, 8] is to systematically eliminate the coefficients which do not convey much information about the image by assigning zero to all the wavelet coefficients at the finer scales corresponding to the same spatial location of a coefficient whose magnitude is below a given threshold. As a result, only a single symbol is transmitted for all the finer scale coefficients. The motivation for using this quantization technique is the assumption that if the energy of a node corresponding to a spatial region of the image is below a given threshold, then the energy of the coefficients at finer scales corresponding to the same region are also expected to be below the threshold. In [4] and [5], the elimination of insignificant subtrees is achieved with respect to different criteria. In [4], the threshold is estimated at every node with respect to the noise sensitivity of the human visual system. This approach reduces the bit rate but introduces distortion if the threshold estimate is not accurate. In [5], a predetermined threshold is applied to a coefficient and to all of its descendents. A single symbol is transmitted for the entire subtree if every coefficient in the subtree is below the threshold.

In this paper, we propose an alternative approach for pyramid image compression using a new zerotree quantizer. The proposed quantization scheme involves a Bayesian decision strategy for successive classifications. A Markov random field progressing from coarse to fine scales [9] is used to improve the prediction of the finer scale coefficients from the coarse scale coefficients.

## 2. PROPOSED ALGORITHM

The first step in the proposed algorithm is to apply a discrete wavelet transform operation to the original image. The wavelet decomposition used in [3] was used for the work presented in this paper. After the wavelet

---

This work was partially supported by the National Science Foundation Graduate Fellowship and by the Advanced Research Projects Agency under contract DABT63-92-C-0022. Address all correspondence to E. J. Delp (ace@ecn.purdue.edu or (317)494-1740)

representation, which consists of three pyramids corresponding to three different spatial orientations, has been obtained, the coefficients in each pyramid are quantized as described below.

With the exception of the lowest frequency decomposition, every wavelet coefficient at a given scale can be related to the coefficient at the previous scale corresponding to the same spatial location and orientation. The coefficient at the coarse scale is known as the parent node. For a given parent, the coefficients of the region at the finer scale corresponding to the same location and orientation are known as children. All the coefficients at all finer scales of the same location and orientation are known as descendants. Zerotree quantizers assign zero to an entire subtree if the subtree is insignificant. In this paper, insignificance is determined using a Bayesian decision scheme.

In order to quantize the coefficients, we would like to classify the individual coefficients into two classes: class 0 and class 1. Following the terminology of [5], a subtree is said to be a zero tree if the entire subtree is assigned to class 0. The coefficient from which the zero tree emerges is said to be a zero root. A coefficient of class 0 which does not belong to a zero tree is said to be an isolated zero. A class 1 coefficient is said to be an isolated one, or a significant coefficient.

Following [5], we assume that the magnitude of the wavelet coefficients at successive scales corresponding to the same spatial location and orientation are correlated. In this paper the correlation is modeled by a multiscale Markov random field model progressing from coarse to fine scale [9]. Let the random field  $Y$  be the pyramid of the absolute value of the wavelet coefficients and the random field  $X$  be the unobserved class labels of the coefficients. At each scale  $m$ , the label field and the observed field are denoted by  $X^{(m)}$  and  $Y^{(m)}$ , respectively. Following [9], it is assumed that the distribution of any random variable in  $X$  given the values of  $X$  at all coarser scales is only dependent on the random variable at its parent node. The random variable  $X_s^{(m)}$ ,  $s \in S^{(m)}$ , where  $S^{(m)}$  is the index set on which  $X^{(m)}$  is defined, can take values from  $\{0, 1\}$ . The dependence of the observed pyramid on its label pyramid is specified through  $f_{Y|X}(y|x)$ , the conditional probability density function of  $Y$  given  $X$ . The conditional distribution of the random variable in  $Y$  at each scale, given its label, is assumed to be Gaussian. The conditional mean and variance of the random variables in  $Y$  from class  $i$ ,  $i = 0, 1$  are denoted  $\mu_i$  and  $\sigma_i^2$ , respectively. The conditional mean  $\mu_i$  is chosen to be the quantization level of the coefficients which belong to class  $i$ . The prior knowledge about the classes is modeled by the probability mass function  $p_X(x)$ . The

transition probability function to model the probability that  $X_s^{(m)}$  is from class  $i$ , given that its parent is from class  $j$ , is chosen as

$$p_{X_s^{(m)}|X_{\partial s}^{(m+1)}}(i|j) = \theta_m \delta(i - j) + (1 - \theta_m) \quad (1)$$

where  $\partial s$  is the index of the parent node of  $X_s^{(m)}$  and the parameter  $\theta_m$  is the probability that the labeling will remain the same from scale  $m + 1$  to scale  $m$ . It is chosen to be an increasing function of resolution.

We would like to assign each pyramid node to one of the four categories: zero root, isolated zero, descendent of zero root, or isolated one. Nodes which are assigned to the categories zero root, isolated zero, or descendent of zero root make up class 0, and the nodes assigned to the category isolated one make up class 1. Beginning at the coarsest scale, each node in the pyramid is assigned to one of the four categories as follows: The likelihood that node  $s$  at scale  $m$  is a zero root is given by the function

$$f_{\hat{Y}_s^{(m)}, \hat{X}_s^{(m)}|X_{\partial s}^{(m+1)}}(\hat{y}_s^{(m)}, \hat{x}_s^{(m)}|\hat{x}) = f_{\hat{Y}_s^{(m)}|\hat{X}_s^{(m)}, X_{\partial s}^{(m+1)}}(\hat{y}_s^{(m)}|\hat{x}_s^{(m)}, \hat{x}) p_{\hat{X}_s^{(m)}|X_{\partial s}^{(m+1)}}(\hat{x}_s^{(m)}|\hat{x}) \quad (2)$$

evaluated with all of the values in  $\hat{x}_s^{(m)}$  equal to zero, where  $\hat{Y}_s^{(m)}$  and  $\hat{X}_s^{(m)}$  are the subtrees in  $Y$  and  $X$ , respectively, rooted at scale  $m$ , node  $s$ , and  $\hat{x}$  is either one or zero, depending on the classification of the parent node of node  $s$ . The first term on the right-hand side of Equation 2 represents the conditional probability density function of  $\hat{Y}_s^{(m)}$  given the classification label subtree  $\hat{X}_s^{(m)}$  and the second term represents the probability mass function of the label subtree  $\hat{X}_s^{(m)}$  given that the parent of node  $s$  has been assigned to class  $\hat{x}$ . Equation 2 is evaluated with all values in  $\hat{x}_s^{(m)}$  equal to zero using the models for  $Y$  and  $X$  described above, and the resulting value is compared to a pre-determined threshold value. If the likelihood of a zero root is greater than the threshold, then node  $s$  is classified as a zero root, and all of its descendants are classified as descendants of zero roots. Otherwise, node  $s$  is classified as an isolated node. To determine if an isolated node belongs to class 0 or class 1, the likelihood function

$$f_{Y_s^{(m)}, X_s^{(m)}|X_{\partial s}^{(m+1)}}(y_s^{(m)}, x_s^{(m)}|\hat{x}) = f_{Y_s^{(m)}|X_s^{(m)}, X_{\partial s}^{(m+1)}}(y_s^{(m)}|x_s^{(m)}, \hat{x}) p_{X_s^{(m)}|X_{\partial s}^{(m+1)}}(x_s^{(m)}|\hat{x}) \quad (3)$$

is maximized over  $x_s^{(m)} = 0, 1$ , i.e., the MAP estimate of  $X_s^{(m)}$  with observed value  $Y_s^{(m)}$  given the class assignment of the parent of node  $s$  is used to decide between isolated zero and isolated one. After each node in

the pyramid has been classified as a zero root, isolated zero, isolated one, or descendent of a zero root, the wavelet coefficients are quantized using the node classifications. Coefficients from class 0, including zero roots, isolated zeros, and zero root descendents are quantized to zero. Coefficients from class 1 are quantized to either  $\mu_1$  or  $-\mu_1$  depending on the sign of the coefficient. After the first round of assignments, the algorithm is repeated to split both class 0 and class 1 into two new finer classes. The process is continued until a target bit rate or desired coded image quality is obtained. Experimental results are presented in the next section.

### 3. EXPERIMENTAL RESULTS

The compression algorithm described in the previous section was applied to the image shown in the upper left of Figure 1. Also shown in Figure 1 are the encoded images obtained for three different data rates. The image in the upper left of the figure is the result from assigning all pyramid nodes at the top level of each of the three pyramids to the category zero root. Thus, all wavelet coefficients except for the low-frequency wavelet coefficients at the top level of the pyramid are quantized to a value of zero. The low-frequency wavelet coefficients at the top level are quantized to 8 bits. The data rate for this image is 0.219 bits/pixel. The image in the lower left of Figure 1 shows the result after one iteration of the compression algorithm described above. The data rate for this image is 0.548 bits/pixel. Finally, the image in the upper right shows the result after two iterations of the compression algorithm. The data rate for this image is 1.409 bits/pixel. It should be noted that for zerotree quantizers, significant reduction in data rates can be achieved by entropy coding the quantized pyramid coefficients [5].

### 4. CONCLUSION

We have presented a new wavelet-based zerotree quantizer which uses a multiscale Markov random field model to model the correlations across scales of the wavelet pyramids. We feel this approach is promising, particularly with respect to progressive transmission of still images.

A postscript version of this paper is available via anonymous ftp to `skynet.ecn.purdue.edu` (Internet address 128.46.154.48) in the directory `/pub/dist/delp/icip94-coding`.

### 5. REFERENCES

[1] I. Daubechies, "Orthonormal bases of compactly

supported wavelets," *Communications on Pure and Applied Mathematics*, vol. 41, pp. 909-996, November 1988.

[2] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674-693, July 1989.

[3] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Transactions on Image Processing*, vol. 1, pp. 205-220, April 1992.

[4] A. S. Lewis and G. Knowles, "Image compression using the 2-d wavelet transform," *IEEE Transactions on Image Processing*, vol. 1, pp. 244-250, April 1992.

[5] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3445-3462, December 1993.

[6] R. A. DeVore, B. Jawerth, and B. J. Lucier, "Image compression through wavelet transform coding," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 719-746, March 1992.

[7] J. W. Woods and S. D. O'Neil, "Subband coding of images," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, pp. 1278-1288, October 1986.

[8] J. M. Shapiro, "Application of the embedded wavelet hierarchical image coder to very low bit rate image coding," *Proceedings of the 1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, April 1993, Minneapolis, Minnesota, pp. 558-561.

[9] C. Bouman and M. Shapiro, "A multiscale random field model for Bayesian image segmentation," *IEEE Transactions on Image Processing*, vol. 3, no. 2, pp. 162-177, March 1994.



Figure 1: Upper left: Original image; Upper right: Image encoded at 0.219 bits/pixel, with all pyramid nodes at top level assigned to zero root; Lower left: Image encoded at 0.548 bits/pixel, after 1 iteration of compression algorithm; Lower right: Image encoded at 1.409 bits/pixel, after 2 iterations of compression algorithm.