Quantization Performance in SPIHT and Related Wavelet Image Compression Algorithms*

Brian A. Banister and Thomas R. Fischer

School of Electrical Engineering and Computer Science

Washington State University

Pullman, WA 99164-2752

e-mail: bbaniste, fischer@eecs.wsu.edu

EDICS No. SPL.IP.1.1

Abstract

The set partitioning in hierarchical trees image coding algorithm is observed to provide progressive classification of the wavelet coefficients. The achievable quantization performance of the induced classes is evaluated for entropy coded scalar quantization and trellis coded quantization, and is compared to the first-order rate distortion function.

1 Introduction

A wavelet decomposed image typically has nonuniform distribution of energy within and across subbands. This provides motivation for partitioning each subband into regions, and assigning each region to a class, based on region energy. This classification approach has led to very effective image compression algorithms, e.g., [1]. Using small classification regions leads to better encoding of each class, but comes at the cost of additional rate for encoding the class maps.

*This work was supported in part by the National Science Foundation under Grant CCR-9627815.

The set partitioning in hierarchical trees (SPIHT) algorithm [2] is an elegant bit-plane encoding method that implicitly classifies the image wavelet coefficients, providing a progressively decodable bit stream. In [3], optimized significance trees are proposed for progressively encoding the coefficient bit planes. However, the underlying scalar quantization of the coefficients is identical to that in the SPIHT algorithm. This letter examines the classification implicit in the SPIHT algorithm, shows that the resulting class data have a multi-mode histogram, and demonstrates that further improvement in quantization performance is achievable.

2 Histogram of classified coefficients

The SPIHT algorithm [2] imposes a hierarchical quad-tree data structure on a wavelet transformed image. Figure 1 indicates the parent-offspring relationship across the subbands. The set of root node and corresponding descendents is referred to as a spatial orientation tree (SOT). Three lists are used in encoding: the list of significant pixels (LSP); the list of insignificant pixels (LIP); and the list of insignificant sets (LIS). The LSP is initialized to be empty, the LIP is initialized with the elements of the lowest frequency subband, and the LIS is initialized with the root of each SOT.

After initialization, a threshold is chosen as $T(0) = 2^{n_0}$, where n_0 is selected such that the largest coefficient magnitude, say c, satisfies $2^{n_0} \leq c < 2^{n_0+1}$. The encoding is progressive in coefficient magnitude, using a sequence of thresholds $T(n) = 2^{n_0-n}, n = 0, 1, \cdots$. Since the thresholds are a power of two, the encoding method can be thought of as "bit-plane" encoding of the wavelet coefficients. At stage n, all coefficients with magnitudes satisfying $T(n) \leq |x| < 2T(n)$ are identified as "significant," and their position and sign bit encoded. This process is called a sorting pass (at stage n). Then, every coefficient with magnitude at least $2T(n)$ is "refined" by encoding the nth most significant bit. This is called a refinement pass. The encoding of significant coefficient position, and the scanning of coefficients for refinement, is efficiently accomplished using the LSP, LIP, and LIS.

A detailed analysis of the SPIHT algorithm reveals that it implicitly uses a form of classification. Whenever a coefficient is found significant at stage n (and moved to the LSP) all insignificant neighbors within a 2×2 coefficient block, and all previously insignificant parent coefficients in the SOT, are also threshold tested and moved to the LIP or LSP. This is effectively assigning all these coefficients (both significant and insignificant) to

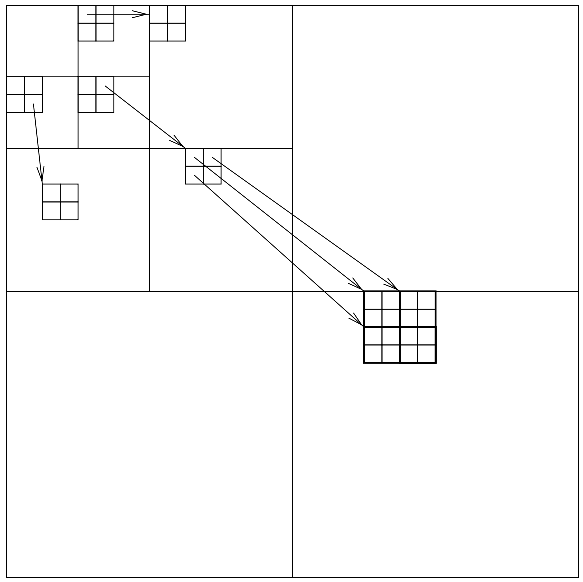

Figure 1: Quad-tree organization of subband coefficients in SPIHT algorithm.

a class, say $C(n)$. In every subsequent stage, all coefficients in the LIP are threshold tested (sorting pass), and hence encoded at one bit per sample (plus a sign bit if found significant). All coefficients in the LSP are refined, and hence also encoded at one bit per sample. Subsequent arithmetic coding can be used to further (although only slightly [2]) reduce the bit rate. The point, however, is that once coefficients are assigned to class $C(n)$, by being placed in either the LSP or LIP during stage n, they are *all* encoded at *every* subsequent stage.

The classification implicit in the SPIHT algorithm is studied by forming a normalized class database. Nine "natural" greyscale images are used to generate a histogram of the classified wavelet coefficients. The coefficients in each class are normalized by dividing each coefficient by the respective class threshold, resulting in normalized coefficients in the range $(-2.0, 2.0)$. Combining the normalized data from each class yields the resulting histogram shown in Figure 2. The histogram is multi-modal and clearly does not match the generalized Gaussian model previously proposed for classification-based subband image coding [1].

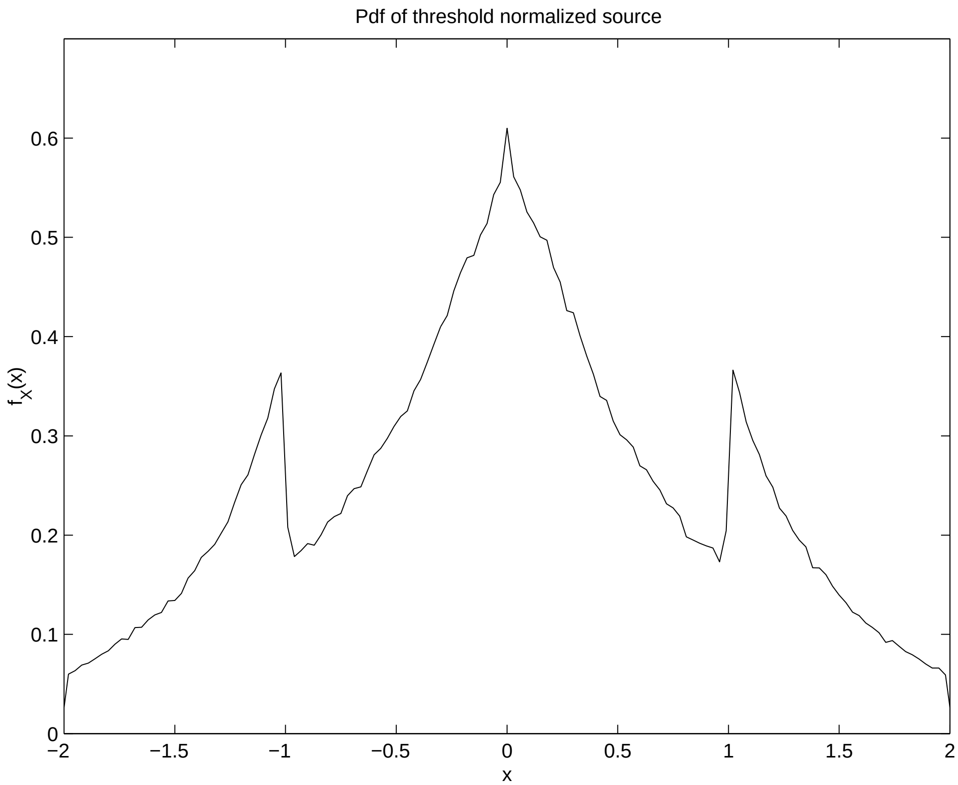

Figure 2: Histogram of threshold normalized coefficients.

3 Quantization Performance

The same normalized class data used to create the histogram is also used to evaluate quantization performance. The mean-squared error (mse) distortions reported are expressed as signal-to-noise ratio (SNR), $10 \log_{10}\left(\sigma^2/mse\right)$, where σ^2 is the variance of the normalized class data.

3.1 Rate-Distortion Function

The Blahut algorithm [4] is used to compute the first-order rate distortion function for the marginal probability density function of Figure 2. This provides a convenient reference for comparison with various quantization methods.

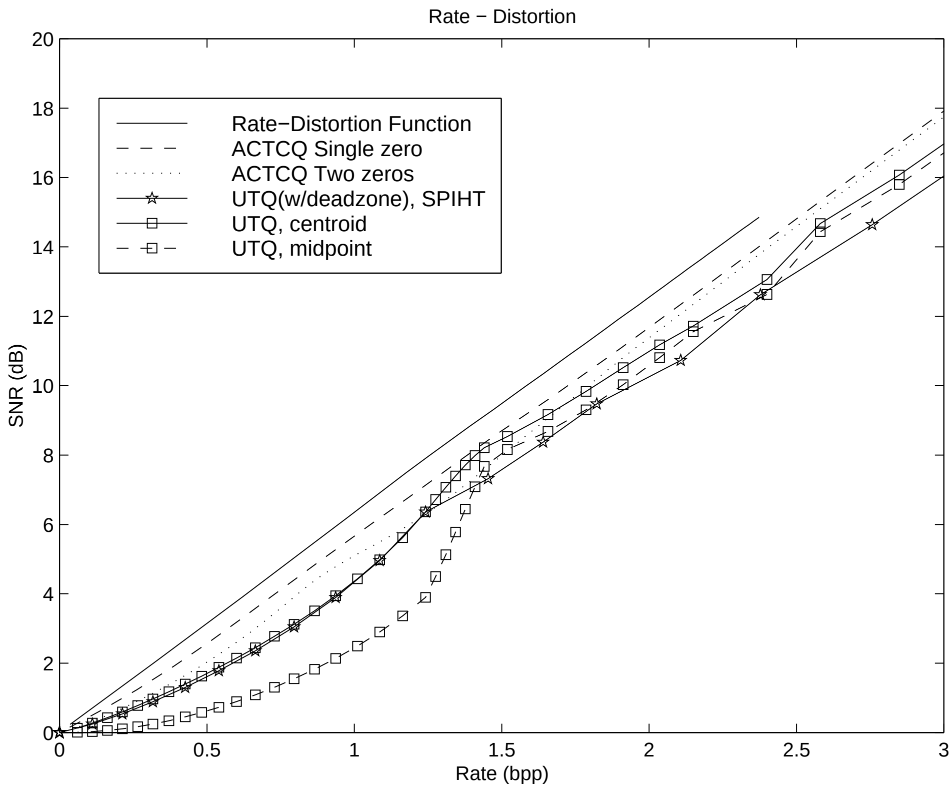

Figure 3: Rate vs. Distortion for several quantizers.

3.2 ECSQ

Entropy coded scalar quantization (ECSQ) performance is evaluated using uniform threshold quantization (UTQ) [5]. Both uniform and uniform-with-deadzone (at the origin) quantizers are used. For uniform quantization, the reproduction levels are either midpoints or centroids. For the quantization with deadzone, the reproduction levels are either midpoints, centroids, or the estimate of centroids used in the SPIHT algorithm. The results are shown in Figure 3 as distortion (SNR) vs. quantizer output entropy.

3.3 ACTCQ

Arithmetic and trellis coded quantization (ACTCQ) [6] is used to encoded the normalized class coefficients. The encoding performance shown in Figure 3 is for a uniform TCQ codebook with either one or two zero codewords. The latter has been found to be appropriate for low-rate encoding of generalized Gaussian sources with small shape parameter [6], and is used in the wavelet/TCQ compression algorithm proposed for JPEG-2000 [7]. Centroids are encoded for the two quantizer regions closest to zero, and midpoints are used for the remaining regions, as in [8]. The ACTCQ performance is based on encoded file sizes, rather than quantizer output entropy as in the ECSQ results.

3.4 Discussion

Figure 3 indicates that by using the SPIHT algorithm to classify wavelet coefficients, further improvement in quantization performance can be obtained. Moreover, close examination of the UTQ results suggests an explanation for the excellent quantization performance achieved by the SPIHT algorithm. Since the number of class coefficients appears to grow roughly exponentially with the classification threshold index [9], at least for relatively small encoding rates, the overall distortion is heavily influenced by the encoding performance for the final class. The SPIHT algorithm classifies in basic 2×2 coefficient blocks, so in addition to the one bit/coefficient comparison required to determine significance, at least one coefficient in four must also have its sign bit coded. If only one coefficient in each block is significant, the average bit rate is 1.25 bits/sample. The UTQ performance curve (using either centroids or the SPIHT reproduction levels) displays a local maximum relative to the rate-distortion curve at an encoding rate of about 1.25 bits/sample. Hence, for that final class encoded at positive rate, the SPIHT UTQ provides especially good performance. It is also interesting to note that the UTQ without deadzone performance is very close to the ACTCQ performance at a rate of about 1.4 bits/sample when centroids are used. Unfortunately, there appears to be no obvious way to easily modify the SPIHT algorithm to use a uniform quantizer without deadzones due to the successive refinement.

Figure 3 also shows that the TCQ codebook with a single zero codeword performs uniformly better than the two zero codebook. This is quite different from the performance for a generalized Gaussian source, and seems to be due to the multi-modal nature of the class density.

References

[1] R. L. Joshi, H. Jafarkhani, J. H. Kasner, T. R. Fischer, N. Farvardin, M. W. Marcellin, and R. H. Bamberger, "Comparison of Different Methods of Classification in Subband Coding of Images," *IEEE Trans. Image Processing*, vol. 6, pp. 1473-1486, Nov. 1997.

[2] A. Said, W. A. Pearlman, "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," *IEEE Trans. Circuits & Systems for Video Technology*, Vol. 6, pp. 243-250, June 1996.

[3] G. M. Davis and S. Chawla, "Image Coding Using Optimized Significance Tree Quantization," *Proc. Data Compression Conf (DCC)*, Snowbird, UT, March 1997.

[4] R. E. Blahut, "Principles and Practice of Information Theory", Addison-Wesley, 1987.

[5] N. Farvardin and J. W. Modestino, "Optimum quantizer performance for a class of non-Gaussian memoryless sources," *IEEE Trans. Inform. Theory*, vol. 30, pp. 485-497, May 1984.

[6] R. L. Joshi, V. J. Crump, and T. R. Fischer, "Image subband coding using arithmetic and trellis coded quantization," *IEEE Trans. Circ. & Syst. Video Tech.*, vol 5, no. 6, pp. 515-523, Dec. 1995.

[7] M. W. Marcellin, "Wavelet/TCQ image compression," Document NCITS/L3.2/97-062.

[8] J. H. Kasner and M. W. Marcellin, "Universal trellis coded quantization," *IEEE International Symposium on Information Theory*, 1995.

[9] W. Zhang and T. R. Fischer, "TCQ subband image coding to exploit dependence," *Conference Proceedings*, 1996 International Conference on Image Processing, Sept. 1996.