Wavelet/TSVQ Image Coding with Segmentation

Keren O. Perlmutter Won Tchoi[†] Sharon M. Perlmutter Pamela C. Cosman[‡]

Information Systems Lab, EE Dept., Stanford University, Stanford, CA 94305-4055
[†]EE Dept., University of Minnesota, Minneapolis, MN 55455
[‡]ECE Dept., University of California, San Diego, La Jolla, CA 92093-0407

Abstract

The use of region-based coding is explored with a wavelet/TSVQ structure. Several methods of generating the segmentation map are discussed, including a recursive segmentation procedure that does not require any side information. The method is investigated on computerized tomographic chest scans, where the images are segmented into three regions – the background, the chest wall region, and the chest organs region. The background is considered of no importance, the chest wall region is considered of low importance, and the chest organs region is considered of high importance. At 0.20 bits per pixel, region-based coding provides a 2.0 dB improvement in the chest organs region at the expense of degradation in the clinically less relevant regions.

1 Introduction

The quality of an image depends on the information content of that image defined with respect to the user's ultimate objectives. In many image compression applications, all areas of an image may not contribute equally useful information for meeting these objectives. For example, medical images typically contain regions that are clinically relevant and regions (e.g., background) that are not. In a videophone image, the head and shoulders are considered to be more vital areas than the background. In such cases, a spatially constant reconstruction quality is not necessary. In addition, in very low bit rate encoding schemes, it may be desirable to ensure good quality for certain areas of an image. Segmentation can address these issues in several ways. It can enable the coder to allocate fewer bits and lower quality to the unimportant regions of an image, and more bits to the relevant areas. Segmentation can also be used to tailor the coding methods for each region. These ideas have been explored by several researchers (e.g., [8, 4, 6, 5]).

The discrete wavelet transform (DWT) combined with scalar or vector quantization has led to a number of effective algorithms for image compression. Using a multiresolution framework, the DWT organizes the coefficients to enable effective quantization and encoding. The DWT has several useful properties that enable it to take advantage of region-based coding. Wavelets have a natural coarse-to-fine representation. This is helpful since this can easily allow good coarse representation for all regions, but good finer representations only for the regions of interest. Another helpful property of wavelets is that each subband maintains a simple spatial relationship to the original image. With each level in the decomposition, the spatial scale of the subbands is reduced by a factor of two in the horizontal direction and a factor of two in the vertical direction. This makes it simple to map the segmentation made in one level into the other levels of the wavelet decomposition.

Pruned tree-structured vector quantization (TSVQ) has a number of useful properties for image compression, as it can provide a variable rate code and has a progressive quality, i.e. each bit improves the quality of the image on the average. A property of pruned TSVQ that makes it particularly amenable to region-based coding is that optimally pruned subtrees are *nested*. Given a large TSVQ, large subtrees (corresponding to higher rate codes) can be used for important regions and smaller subtrees (corresponding to lower rate codes) can be used for unimportant regions. Since these trees are nested, memory storage is reduced, as the smaller tree is simply a subset of the larger tree.

In this paper, we explore the use of segmentation in the refinement of the *bit allocation* procedure used in the design of pruned TSVQs for wavelet coefficient quantization.

2 Algorithm

In the design of our system, the Daubechies orthogonal 8-tap filter is used to decompose the training images 4 levels. This decomposition produces 13 subbands. The wavelet decomposition is illustrated in Figure 1. The next step involves the construction of separate codebooks for each subband. The lowest band is scalar quantized, the finest scales are encoded using 4 $\times$4 vectors and intermediate scales employ vectors of size 2, 4, and 8. A large TSVQ is greedily grown on the training sequence associated with each band using the generalized Lloyd algorithm and is then optimally pruned back [1]. The final stage of the design provides the bit allocation for each vector, which is determined by two factors: the subband in which the vector is located and the segmentation region in which it lies. The first factor is based on the fact that fewer bits can be allocated to the high frequency subbands than to the low frequency ones since they possess smaller variance, corresponding to less information. The bit allocation can be performed by selecting appropriate subtrees for each band. For each band, $i \in \{1, \ldots 13\}$, the distortion and rates for the sequence of pruned subtrees provide distortion-rate $D_i(R_i)$ curves. The minimum overall distortion for a given total rate can be achieved by selecting points of equal slope across the individual distortion rate curves of the different subbands [7]. The second factor refines this bit allocation by selecting different points of equal slope based on the segmentation map. For important regions, a low value for the point of equal slope will be selected, thus generating a set of relatively large subtrees. On the other hand, at less important regions a high value for the point of equal slope is selected, thereby generating a set of relatively small subtrees.

3 Segmentation Map Information

Given a segmentation for the original image, there are several alternatives for generating segmentation maps for each of the subbands. The construction of the maps depends on whether side information indicating the region information can be tolerated in the system. If side information is acceptable, the map can be generated on the original image, and then, because of the spatial relationships that exist between the bands, the maps corresponding to the subbands of the different levels can be obtained by downsampling the original map. (Note that with all of the segmentation methods presented in this work no map is used for subband 1). Because the decoder can also perform the downsampling operation, only the highest resolution map must be transmitted to the decoder. This side information can be losslessly compressed.

Side information may not be desirable in some situations, however. For example, at very low rates, it may not be desirable to send any overhead to the decoder to

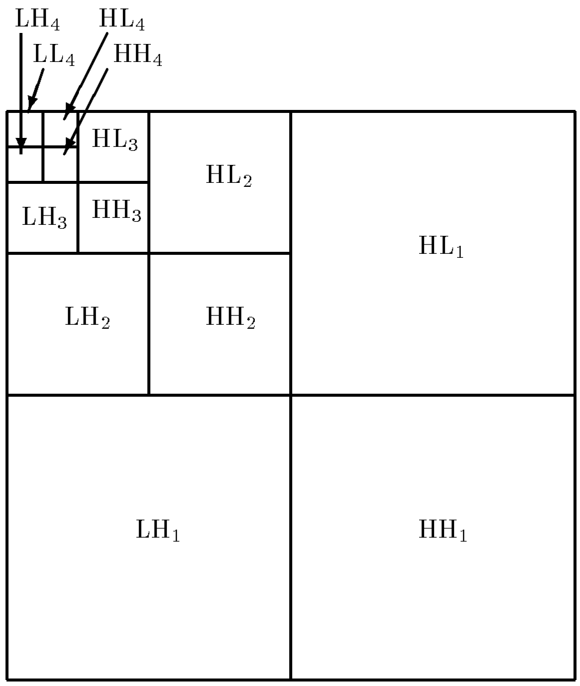

Figure 1: Wavelet decomposition

relay the segmentation information. In addition, complicated images may lead to costly descriptions of the segmentation map. Furthermore, channel errors might corrupt side information, which may have a great impact on the reconstructed image if large portions of the important region are misclassified as unimportant. Two different methods can be used to generate maps with no side information. Both methods begin by encoding the baseband (subband 1) with the TSVQ associated with that subband. Then, the encoded baseband is segmented. One method for producing the segmentation maps for the upper level bands is to upsample the segmented baseband. The idea of avoiding side information by examining the baseband was explored by Johnsen et al. in [2]. In that work, the edge location information was upsampled to indicate edge locations in the upper bands. In the second method, rather than simply upsampling, the segmentation map of the upper level subbands can be refined at the expense of additional complexity. A recursive segmentation, based on Mohsenian's structure introduced for edge location identification [3], can be used to refine the upper band segmentation without the use of any side information. This procedure is illustrated in Figure 2, and proceeds as follows: The lowest band (LL_4) of size 32×32 is encoded with Lloyd-Max scalar quantization. The reconstructed version of this band is

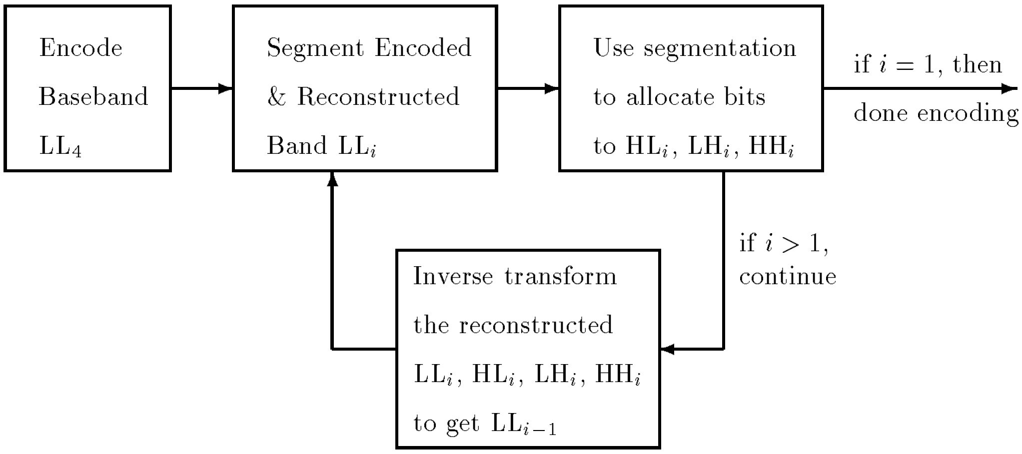

Figure 2: The encoder structure involves recursively segmenting, quantizing, and inverse transforming

segmented, and the segmentation guides the encoding of the 2-D vectors in the remaining 3 subbands of size 32×32, (HL_4, LH_4, and HH_4). Within each band, different regions use different subtrees. The decoder is able to reconstruct the LL_4 band and apply the same segmentation, thereby determining the bit allocation to follow. After subbands HL_4, LH_4, and HH_4 are encoded and transmitted to the decoder, both encoder and decoder can reconstruct these bands and combine them with the reconstructed LL_4. Performing one stage of a wavelet inverse transform then yields a new baseband, LL_3. The encoder and decoder then segment that baseband, and allocate bits for the 4-D vectors of bands HL_3, LH_3, and HH_3. This recursive structure continues for the LL_2 and LL_1 basebands.

One other issue must be addressed in determining the segmentation map. Because vectors are classified, rather than pixels, there is the possibility that pixels within a particular vector possess different region affiliations. The vector classification can be resolved by majority rule, by choosing the region of highest importance, or by choosing the region of least importance.

4 CT Segmentation

In this work we investigate the use of region-based segmentation on a set of computerized tomographic (CT) chest scans. An original 11-bit CT chest scan (printed with 8-bit grayscale) is shown in Figure 3(a). Figure 3(b) shows the segmentation of the original image into 3 regions. The chest organs are shown in white. This region contains the lungs and the mediastinum (the portion of the chest which has the heart and the great blood vessels). The chest wall region, comprising the muscles and ribs, is shown in gray. The background of the image, including the table on which the patient lies, is shown in black. This segmentation can be achieved automatically for many CT chest scans using a simple amplitude threshold and pixel adjacency rule. Because the background is considered of little or no importance, it is assigned zero bits (except in subband 1). The chest wall region is encoded at a low rate (using small subtrees), and the chest organs region is encoded at a high rate (using large subtrees).

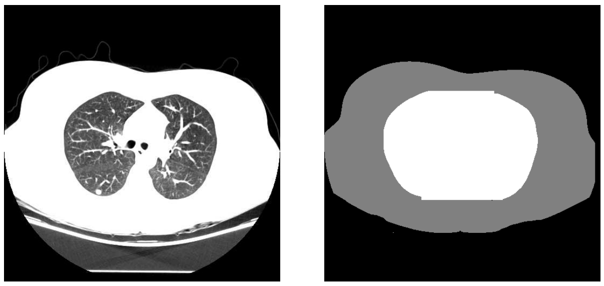

Figure 3: Original and segmented CT chest scans

5 Results and Discussion

The training sequence was composed of 10 CT chest scans. Each image was transformed, and the various bands blocked into vectors of the size and shape chosen for that band. A CT chest scan from a different patient study was used as a test image. We deemed the chest organs region to be 25 times more important than the chest wall region, and thereby assigned to the former region a slope 25 times lower during the bit allocation procedure. The classification of each vector was based on the most important region contained within the vector.

Figure 4 displays a comparison of the SNR results obtained on the chest organs using wavelet/TSVQ coding without segmentation and the wavelet/TSVQ coding with the recursive segmentation procedure. Here SNR is defined as $\text{SNR} = 10\log_{10}(D_0/D)$, where D is the distortion measured by mean squared error and D_0 is the distortion produced by the best zero rate code. The x-axis shows the average bit rate for the entire image. The graph thus illustrates the differences in SNRs obtainable by the two methods for that region when the entire encoded images are constrained to have the same bit rate. The figure shows that coding with segmentation provides up to 2.0 dB improvement in the chest organs region compared to coding without segmentation. A graph of the SNRs for the chest wall region would show a similar difference, except within this region, the segmentation method would produce the inferior performance. This is illustrated in Table 1, which provides results for the two methods for the three regions at three different rates. For example, the table illustrates that at 0.20 bpp the increase in SNR of 2.0 dB for the region of clinically important chest organs comes at the expense of greatly increased noise in the background, which is of no clinical relevance at all, and a 1.5 dB loss in the chest wall region, which is rarely of clinical importance. The background loss comes partly from the severe blurring of the table on which the patient lies. Figures 5(a) and (b) show the encoded images with recursive segmentation and without segmentation. Both images have been windowed at intensities 50 and 400 and then linearly rescaled for 8-bit printing. This windowing displays most clearly the degradation that the segmentation causes in the background and chest wall regions. Figures 5(c) and (d) show a portion of the lungs of the same encoded images, also windowed at 50 and 400, which demonstrate the higher SNR and better visual quality in the lung region produced by the segmentation method.

Comparison of Segmentation Maps

The recursive procedure improved the segmentation of the image considerably over the simple upsampling method. In particular, Figure 6(a) illustrates the map generated at 0.20 bpp using the upsampling method and Figure 6(b) illustrates the map generated at the same rate using the recursive method. Table 2 indicates the classification errors produced by the maps compared to the downsampled version of the original map (this is considered the "true" map), which illustrates the classification improvement gained by the recursive procedure. The improved segmentation produced by the recursive procedure led to small improvement (0.2-0.3 dB) in the chest organs region and slightly more improvement (up to 0.8 dB) in the chest wall region.

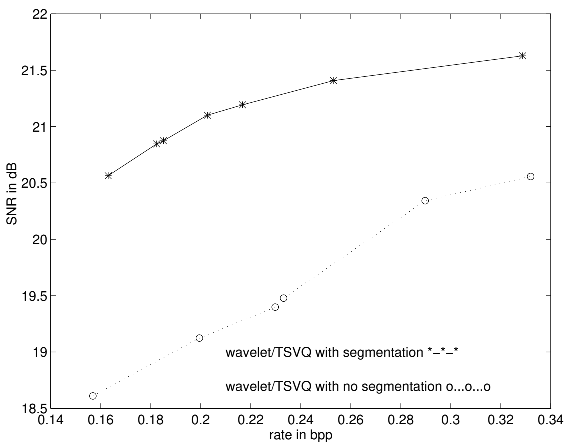

Figure 4: SNRs achievable for chest organs region

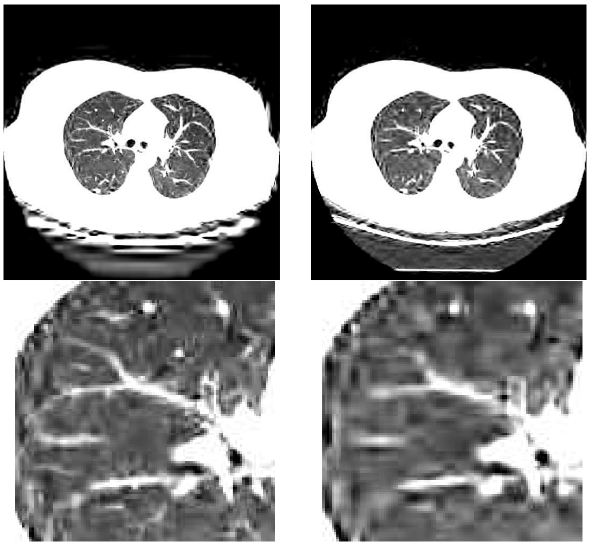

Figure 5: Top row: Encoded images at 0.20 bpp (a) wavelet/TSVQ with recursive segmentation, (b) wavelet/TSVQ with no segmentation. Bottom row: Zoom of encoded lung regions (c) wavelet/TSVQ with recursive segmentation, (d) wavelet/TSVQ with no segmentation

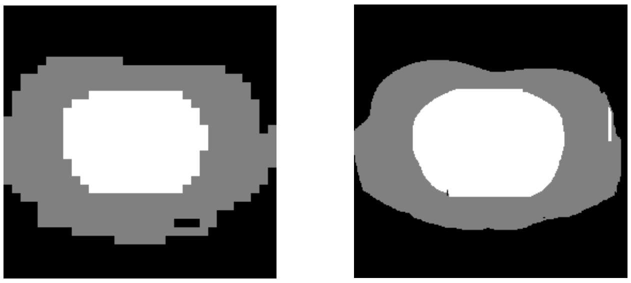

Figure 6: Maps of size 256x256 for CT chest scans at 0.2 bpp using (a) the upsampling method and (b) the recursive segmentation procedure

If side information for the map description can be tolerated, we can also compare the recursive segmentation procedure to the downsampling method that uses overhead information to characterize the highest resolution map. For the CT example we consider here, the image can be easily described. As a result, side information does not contribute much overhead to the decoder. Consequently, the two methods provide comparable performance for the chest organs region.

6 Conclusions

In this paper, region-based segmentation in conjunction with wavelet/TSVQ coding was investigated. The algorithm was able to take advantage of several properties of wavelets and TSVQs in allocating bits to the different regions. Several segmentation map generation schemes were described, including a recursive procedure that did not require any side information to transmit the segmentation information. On a CT image, region-based coding led to a 2.0 dB improvement in the clinically important region compared to a method that did not segment. Recursive segmentation produced a refined segmentation map compared to simply upsampling the baseband map. For this easily described CT image, sending the actual maps was not very costly and produced comparable results in the most important region. We note however that in some situations, such as very low bit rate coding and the transmission of multi-region and complicated images, side information may not be desirable. In such cases, recursive segmentation may be a viable option to use rather than simply transmitting the segmentation information to the decoder.

Acknowledgments

This work was funded by the National Science Foundation under grant NSF MIP-9016974, an NSF postdoctoral fellowship, and an AT&T fellowship.

References

[1] A. Gersho and R. M. Gray. *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, Boston, 1992.

Region	Recursive Segmentation	Upsampling
Background	0.5	11.9
Chest Wall Region	2.1	4.2
Chest Organs Region	1.3	9.9
Total	1.1	9.2

Table 2: Misclassification Percentages

[2] O. Johnsen, O.V. Shentov, and S.K. Mitra. A technique for the efficient coding of the upper bands in subband coding of images. In *Proceedings ICASSP*, pages 2097–2100, Albuquerque, April 1990.

[3] N. Mohsenian and N. Nasrabadi. Edge-based subband VQtechniques for images and video. *IEEE Trans. on Circuits and Systems for Video Technology*, 4(1):53–67, Feb. 1994.

[4] E. Nguyen, C. Labit, and J-M. Odobez. A ROI apporach for hybrid image sequence coding. In *IEEE International Conference on Image Processing*, volume 3, pages 245–249, Austin, TX, 1994.

[5] S. M. Perlmutter, K. O. Perlmutter, P. C. Cosman, E. A. Riskin, R. A. Olshen, and R. M. Gray. Tree-structured vector quantization with region-based classification. In *Proc. Twenty-sixth Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Oct. 1992.

[6] G. Poggi and R. A. Olshen. Pruned tree-structured vector quantization of medical images with segmentation and improved prediction. *IEEE Trans. on Image Processing*, 4(6):734–742, 1995.

[7] T. Senoo and B. Girod. Vector quantization for entropy coding of image subbands. *IEEE Trans. Image Process.*, 1(4):526–532, Oct. 1992.

[8] H. D. Fisher T. W. Ryan, L. D. Sanders. Wavelet-domain texture modeling for image compression. In *IEEE International Conference on Image Processing*, volume 2, pages 380–384, Austin, TX, 1994.

Method	Background	Chest Wall	Chest Organs
Wavelet/TSVQ with no segmentation: 0.16 bpp	13.6	25.2	18.7
Wavelet/TSVQ with recursive segmentation: 0.16 bpp	5.3	23.5	20.6
Wavelet/TSVQ with no segmentation: 0.20 bpp	14.2	25.8	19.1
Wavelet/TSVQ with recursive segmentation: 0.20 bpp	5.3	24.3	21.1
Wavelet/TSVQ with no segmentation: 0.23 bpp	14.5	26.3	19.4
Wavelet/TSVQ with recursive segmentation: 0.23 bpp	5.4	24.6	21.2

Table 1: SNR results for all 3 regions of CT image