TRELLIS-CODED RESIDUAL VECTOR QUANTIZATION BASED PYRAMID IMAGE CODING

Mohammad A. U. Khan

Department of Electrical Engineering
King Fahd University
of Petroleum and Minerals
Dhahran 31261, Saudi Arabia
maukhan@kfupm.edu.sa

ABSTRACT

A recent algorithm for residual vector quantization is used for coding image subbands. The algorithm, called trellis-coded residual vector quantization, designs expanded codebooks by using trellis-coded quantization technique. The new algorithm also takes advantage of multiple-stage quantization feature embedded in a residual vector quantization structure. The union of residual vector quantization (RVQ) and trellis-coded vector quantization (TCVQ) in trellis-coded residual vector quantization provides improved rate-distortion performance for image subbands. Issues regarding the construction of image subbands and the bit allocation problem are discussed in the paper. Simulation results for test images are presented and compared with other trellis-coded quantization based subband coding techniques.

1. INTRODUCTION

Subband coding was introduced by Crochiere *et al.* in 1976 for speech. Since then, considerable attention has been devoted to subband coding of speech and it has proven to be a powerful technique for the medium and low bandwidth speech coding. In 1986, Wood and O'Neill demonstrated the effectiveness of subband coding technique for images. The basic idea of subband coding is to split up the frequency band into subbands and quantized and transmitted to the deocder. In a subband decomposed image, the different subbands usually contain vastly different amounts of energy. This property of subbands is utilized in coding. The bands which contain more energy are quantized using a finer quantizer and those bands which contain less energy are quantized more coarsely.

Two important issues in subband image coding are the selection of a quantization scheme and the distribution of bits among various bands. The encoding of each subband is usually carried out for each subband separately. However, it is possible to jointly encode the subbands to take advantage of inter-subband dependencies [1]. A scheme that is more adaptive to image statistics is relatively easy to design by separately encoding the subbands. Furthermore, separately encoding the subbands makes it possible for us to see the effect of quantization more clearly. Therefore, in this paper we opted for separately encoding the subbands.

Most of the energy within subbands is confined to areas corresponding to edges and strong textures in the original image. The bands can be easily divided in two categories: one corresponding to low frequencies present in the image and the other containing high frequency details. The bands which belongs to the first category exhibit strong two-dimensional correlation, while the other category has bands with very small amount of correlation. In this paper, we proposed to apply conditional entropy-constrained trellis-coded residual vector quantization (CEC-TCRVQ) for the first category bands and entropy-constrained trellis-coded residual vector quantization (EC-TCRVQ) for the second category bands.

Conditional entropy-constrained trellis-coded residual vector quantization (CEC-TCRVQ) is an extension to the entropy-constained trellis-coded residual vector quantization (EC-TCRVQ), introduced in [2]. The CEC-TCRVQ employ adjacent vector conditioning to take advantage of the two-dimensional correlation present in a low-frequency band.

The organization of this paper is as follows: Section 2 discusses the details of subband decomposition with an octave band structure. Designing of codebooks for various subbands is described in Section 3. Conditioning structure to exploit the intra-band correlation is mentioned in Section 4. Section 5 describes the bit allocation strategy in detail. Section 6 presents the simulation results and comparison with a trellis-based subband image coding technique.

2. CONSTRUCTION OF IMAGE SUBBANDS

In our proposed scheme, the image is split into a pyramid, and each pyramid subband is coded independently. The pyramid construction process begins with the splitting of the image into four subbands, and then continues with the division of the lowpass band recursively up to the required level of decomposition. Here we use three levels of decomposition to get ten subbands. This spectral decomposition is shown in Figure 1.

To get the image pyramids, we use 9-7 bi-orthogonal filters which are popular in the wavelet image coding community. Partly because many state-of-the-art image compression methods, like EZW [3], LZC [4], MTWC [5], and SPIHT [6], use this wavelet transform as their first step.

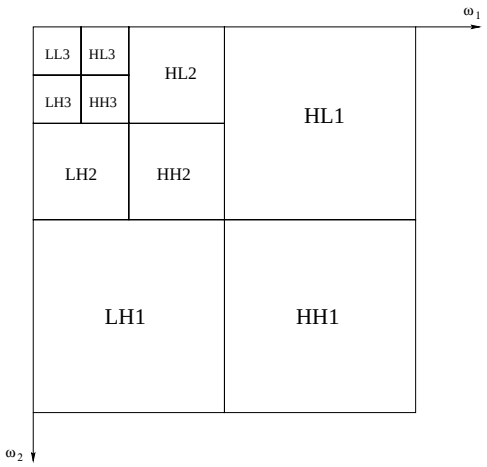

Figure 1: Image decomposition

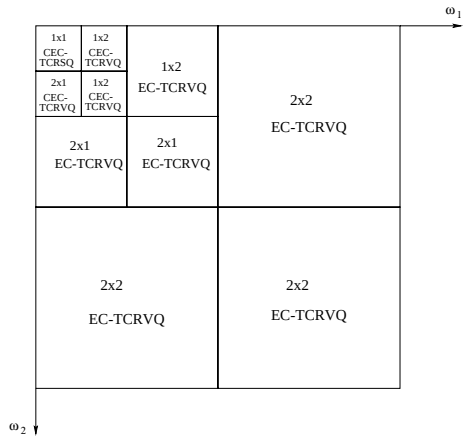

Figure 2: Trellis-coded quantization scheme for image subbands

The coefficients of the right half of the analysis bank are given in Table 1. The left side of the filter is generated by making use of the symmetry of the filters. The synthesis bank is generated using the following relationships:

$$\begin{aligned} \tilde{g}(n) &= (-1)^{1-n}\, h(1-n), \quad \text{and} \\ \tilde{h}(n) &= (-1)^{n}\, g(1-n). \end{aligned} \tag{1}$$

Table 1: 9-7 Biorthogonal Filter Banks

Filter type	Right-side Filter Coefficients	Filter Length
Low pass	$\frac{1}{\sqrt{2}} \times [0.853, 0.3774, -0.11062, -0.023849, 0.03783]$	9
High pass	$\frac{1}{\sqrt{2}} \times [0.7885, -0.4181, -0.04069, 0.06454]$	7

3. DESIGNING OF CODEBOOKS

We used trellis-coded residual vector quantization for designing the codebooks of the image pyramids by employing a training set of 14 (512×512) images. The image subbands usually differed in their spectral contents [7], therefore normalized codebooks were designed for each subband by dividing all of the training data by their respective standard deviations. The mean of the baseband (LL3 band) was also subtracted. Therefore, the mean of the baseband and the standard deviations of its ten bands needed to be sent to the decoder. This overhead information corresponds to a negligible increase in the overall bit rate.

Figure 2 shows various trellis-coded residual vector quantization schemes used to quantize the subbands. The LL3 band which contains the texture, also contains strong two-dimensional correlation. In order to effectively exploit the correlation to reduce the bit rate, we coded the LL3 band using three-stage conditional entropy-constrained trellis-coded residual scalar quantization (CEC-TCRSQ). The reason for using a scalar quantizer lies in the fact that it is difficult to code textures using vector quantization without producing visual artifacts. The bands HL3, HH3 and LH3, also contain some vertical and horizontal correlation so we employed three-stage two-dimensional conditional entropy-constrained trellis-coded residual vector quantization. There is little correlation present in the HL2, HH2, and LH2 bands. Therefore, we used four-dimensional trellis-coded residual vector quantization for these bands. The HL1, HH1 and LH1 bands contained very small correlation and also a small amount of image energy. Hence we needed to code them at low bit rates. In our scheme, we coded these bands using 16-dimensional trellis-coded residual vector quantization.

4. THE USE OF CONDITIONAL ENTROPY-CONSTRAINED QUANTIZATION IN SUBBAND IMAGE CODING

After the subband decomposition of an image, dependencies may exist among the coefficents of the same band, called intra-band dependencies and also across the bands, referred to as inter-band dependencies. Although the dependencies that still exist after subband decomposition are not as strong as they were in the spatial domain, nevertheless the coding performance can be improved by employing quantization schemes that exploit those dependencies. For that purpose, differential pulse code modulation (DPCM) was applied successfully to all the subbands in [8] and to the low-pass subband in [9] and [10].

In [1], authors tried to use Lagrangian formulation in order to jointly optimize the subband coder. The joint optimization framework was constructed by viewing the set of subband quantizers in terms of a structurally constrained *product* quantizer, consisting of all possible combinations of outputs from subband quantizers. The optimality conditions are developed and an iterative design algorithm can be formulated that satisfies these conditions. The conditioning support for this technique is spread among the subbands. The best inter- and intra-band conditioning symbols are s-

elected among vectors closer in space or frequency to the current vector. This scheme has resulted in improved performance over all other subband coding schemes.

We also used the Lagrangian formulation for exploiting the correlation present in a given subband. For this purpose, we used conditional entropy-constrained trellis-coded residual vector quantization for image bands that exhibit large correlation. We based the intra-stage conditioning in a specific band based on its orientation. For the LL3 band in Figure 8.1 we chose the left most vector and the vector on top of the given input vector for the second-order inter-stage conditioning. The HL3 band, which contained horizontal correlation, was coded using inter-stage conditioning of the left most vector. The LH3 band makes use of a vertical top vector for inter-stage conditioning. The rest of the bands are coded using entropy-constrained trellis-coded RVQ, which does not employ any inter-stage conditioning.

5. BIT ALLOCATION

Once the quantization scheme is specified for the image pyramids, the next issue is how to distribute the bit budget among the subbands. Westerink, Biemond, and Boekee [11] developed an optimal bit allocation algorithm based on the subband variance. Riskin [12] restated their algorithm using the generalized BFOS algorithm for both cases of convex and non-convex operational distortion-rate functions.

Bit allocation using the generalized BFOS algorithm was first suggested in [13]. The generalized BFOS algorithm is an extension of an algorithm for optimal pruning, in tree-structured classification and regression [14], to coding. For a source coding application, it finds a sequence of nested subtrees of a given tree-structured coder. The selection of a subtree is optimal in that it has the lower average distortion of all subtrees of the tree with the same or lower average rate.

In our context of bit-allocation, the BFOS algorithm can be used as follows: construct a tree T with l subtrees where each subtree is a unary tree and represents a subband. In each subtree we have k nodes where each node is represented by an (R, D) point found during the quantization design. The topmost node in the subtree corresponds to the zero-rate codebook, i.e., the Lagrangian parameter λ has its maximum value. Moving down in the subtree, we decrease the λ values, and at the bottom of this subtree, we find the (R, D) point that corresponds to $\lambda = 0$. This node is called "the leaf node". According to this organization, if we move downward along the subtree, the rate at each node increases, and the distortion decreases. Besides the (R, D) point stored at each node, we also store a parameter called s which is defined to be the ratio $\delta D/\delta R$, where δD and δR denote the magnitude of the difference between the distortion and the rate at the current node and the leaf node respectively. The value s can also be interpreted as a slope that trades off D and R in each subband.

If we denote the initial tree by T_I, the generalized BFOS algorithm will prune off the branches of the initial tree in order to form the final pruned tree T_F. In this pruning operation, the algorithm obtains a sequence of trees where each intermediate tree T_{i+1} is obtained by pruning off the node having the smallest slope s in the tree T_i. The pruned leaf node belongs to a certain subtree, and therefore this iteration provides a new leaf node in the previous tree. After this procedure, the s ratio must be re-calculated in this new tree T_{i+1}. The algorithm ends when the sum of the leaf node rates drops below the target rate. The codebook used to encode each subband corresponds to the codebook specified by the leaf nodes of the final pruned tree T_F.

6. SIMULATION RESULTS

In this section, we present results for 512×512 Lena image at low bit rates. We compared our results with quad-tree based trellis-coded quantization (QTCQ) recently developed by Banister and Fisher [15],[16]. QTCQ employs a form of wavelet coefficient classification used in the SPIHT algorithm [6]. This classification is based on the quad-tree structure and moves across the subbands to get classed with coefficients with minimum intra-class correlation. Then these magnitude-threshold-based classes are coded using arithmetic-coded trellis-coded scalar quantization. In order to provide a fair comparison, we also used a three-level decomposition in the QTCQ.

For the bit allocation tree, we obtained thirty rate-distortion pairs for each subband. The tree has ten branches with thirty points on each branch. The bit allocation algorithm described earlier in section 8.4 is applied to the tree. Table 6 shows the final results among the various subbands of Lena for the rate of 0.25 bits per pixel. The *total distortion* value in the table is obtained by adding the individual distortion of all the subbands. The *actual distortion*, listed at the bottom, is the actual mean squared-error between the coded and original images. This value differs from the total distortion because the filters are not ideal. Another interesting point is that the mismatch between the actual result and the result inside the training set (TS) is low. This happens because we use normalized codebooks.

Subband	Training set		Lena	
	D	R	D	R
LH1	2.137	0.001	3.930	0.013
HL1	6.542	0.103	5.589	0.123
HH1	3.446	0.006	2.094	0.005
LH2	4.683	0.121	3.565	0.175
HL2	6.769	0.451	5.554	0.468
HH2	2.667	0.428	2.379	0.439
LH3	0.985	1.374	0.959	1.284
HL3	1.204	1.983	1.248	1.855
HH3	0.964	1.484	0.462	1.233
LL3	0.954	6.013	0.868	5.455
Total	30.35	0.27	26.648	0.25
Actual Distortion = 28.169				

Table 2: Bit allocation for test image Lena at 0.25 bits per pixel.

We also coded image Moffett from University of California Los Angles image database. The results are shown in Figure 6.

Figure 4 compares our TCRVQ-based subband coder (TCRVQ-SBC) with other results in the literature for the

(a) (b)

Figure 3: Image Moffett coded at 0.22 bits per pixel using (a) QTCQ (b) our proposed image coder.

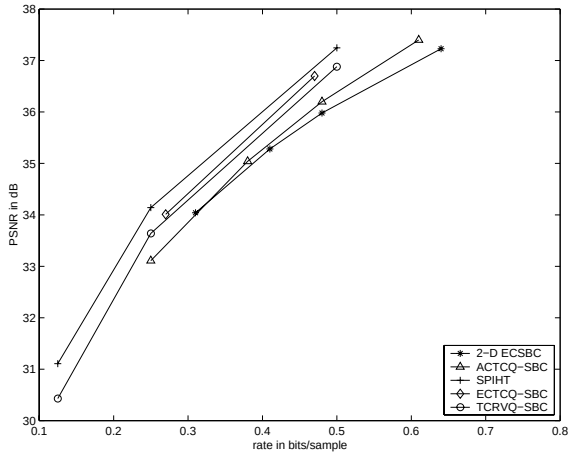

Figure 4: Comparison of subband coding performance for "Lena".

test image Lena. Kim and Modestino [17] report PSNR's of 34.04 dB, 35.28 dB, 35.98 dB, 37.23 dB for bit rates of 0.31, 0.41, 0.48, and 0.64 bpp, respectively, for their entropy-constrained subband coder (2-D ECSBC). Joshi, Crump and Fischer [18] devloped arithmetic-coded trellis-coded subband image coder (ACTCQ-SBC) and is shown to provide about 0.25 dB improvement over the 2-D ECSBC design. Sriram and Marcellin [19] report PSNR's of 34.01, 36.70, and 40.06 dB for bit rates of 0.27, 0.47, and 0.95 bits per pixel, respectively, for their entropy-constrained trellis-coded quantization based subband image coder (ECTCQ-SBC). SPIHT [6] results are also displayed in the figure. The figure shows that our coder does better than the ACTCQ-SBC and the 2-D ECSBC. Comparing the performance of our coder with that of ECTCQ-SBC shows that TCRVQ-SBC performance is worse by about 0.5 dB at 0.5 bits per pixel and is about 0.15 dB worse at 0.25 bits per pixel. This may be due to the reason that ECTCQ is a single stage system as compared to ECTCRVQ. The TCRVQ-SBC performs worse in comparison to SPIHT by about 0.6 dB. We believe that this gap is due to the reason that SPIHT coder exploits inter-band dependence while our codeer does not.

7. CONCLUDING REMARKS

Our proposed subband-coder seems to provide low-contrast images at low bit rates compared to the QTCQ-based scheme. On the other hand, our proposed scheme does an excellent job at preserving low-magnitude textures present in some images. The QTCQ-based scheme which employs magnitude-based classes does not make it to low magnitude classes as its bit budget is finished well before that, and hence does a poor job at preserving the small-magnitude texture.

8. REFERENCES

[1] F. Kossentini, W. Chung, and M. Smith, "Subband image coding with jointly optimized quantizers," in *IEEE Int. Conf. Acous., Speech and Signal Processing*, (Detroit, MI), pp. 2221–2224, April 1995.

[2] M. A. Khan, M. J. Smith, and S. W. McLaughlin, "Cinditional entropy-constrained trellis-coded rvq

with application to image coding," *IEEE Trans. on signal processing letters*, vol. 7, pp. 49–51, March 2000.

[3] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Signal Processing*, vol. 41, pp. 3445–3462, Dec. 1993.

[4] D. Taubman and A. Zakhor, "Multirate 3-d subband coding of video," *IEEE Trans. on Signal Processing*, vol. 3, pp. 572–588, Sept. 1994.

[5] H.-J. Wang and C.-C. J. Kuo, "A multi-threshold wavelet coder (mtwc) for high fidelity image," in *International Conference on image processing*, July 1997.

[6] A. Said and W. 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, pp. 243–250, June 1996.

[7] B. Mahesh and W. Pearlman, "Multiple-rate structured vector quantization of image pyramids," *Journal of visual communications and image representation*, vol. 2, pp. 103–113, June 1991.

[8] J. Woods and S. O'Neil, "Subband coding of images," *IEEE trans. on Acoustic, Speech and signal processing*, vol. ASSP-34, pp. 1278–1288, october 1986.

[9] H. Gharavi and A. Tabatabai, "Sub-band coding of digital images using two-dimensional quadrature mirror filtering," in *Proc. of SPIE, Visual communications and image processing*, vol. 707, pp. 51–61, 1986.

[10] P. Westerink, J. Biemond, and D. Boekee, "Evaluationof image subband coding schemes," in *Eur. Signal Proc. Conf. (EUSIPCO)*, (Grenoble, France), pp. 1149–1152, September 1988.

[11] P. Wasterink, J. Biemond, and D. Boekee, "An optimal bit allocatiohn algorithm for sub-band coding," in *proceedings of ICASSP*, pp. 757–760, 1988.

[12] E. Riskin, "Optimal bit allocation via the generalized bfos algorithm," *IEEE Trans. on Information Theory*, vol. 37, pp. 400–402, Mar 1991.

[13] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and regression trees.* Belmont, California: The wadsworth statistics/probability series, Wadsworth, 1984.

[14] P. Chou, T. Lookabaugh, and R. Gray, "Optimal pruning with applications to tree-structured source coding and modeling," *IEEE trans. on information theory*, vol. 35, pp. 229–315, March 1989.

[15] B. Banister and T. Fischer, "Analysis of classification and quantization in spiht and related wavelet image compression," in *submitted to IEEE signal processing letters*, March 1998.

[16] B. Banister, "Quad-tree based classification in image compression," Master's thesis, Washington state university, August 1998.

[17] Y. Kim and J. Modestino, "Adaptive entropy-coded subband coding of images," *IEEE trans. on image processing*, vol. 1, pp. 31–48, Jan 1992.

[18] R. Joshi, V. Crump, and T. Fischer, "Image subband coding using arithmetic-coded trellis-coded quantization," *IEEE Trans. on circuits and systems for video technology*, vol. 5, pp. 515–523, Dec. 1995.

[19] P. Sriram and M. Marcellin, "Image coding using wavelet transforms and entropy-constrained trellis-coded quantization," *IEEE Trans. on Image Processing*, vol. 4, pp. 725–733, June 1995.