FAST IMAGE CODING USING BLOCKWISE BINARY CLASSIFICATION AND TRELLIS CODED QUANTIZATION

Zhihai He, Tian-Hu Yu, and Sanjit. K. Mitra

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106
{zhihai, mitra}@iplab.ece.ucsb.edu

ABSTRACT

With blockwise binary classification and data partitioning, we convert the image subbands to the type of data for which the trellis coded quantization (TCQ) has the best quantization performance. Compared to the arithmetic coded TCQ (ACTCQ)[3] and other TCQ-based coding schemes, the proposed algorithm reduces the computational complexity by about 90%. However, it performs competitively with the best available coding algorithms reported in the literature in terms rate-distortion performance.

1. INTRODUCTION

Trellis coded quantization (TCQ), introduced by Marcellin and Fischer [1, 2], has excellent rate-distortion (R-D) performance for the memoryless uniform or generalized Gaussian source. Several methods have been reported in the literature that apply the TCQ to wavelet image coding.

In the ACTCQ coding algorithm [3], the input image is first decomposed into 16 subbands with the lowest frequency subband further transformed by a 4×4 DCT. Each subband is then modeled as a memoryless Gaussian source. Based on the pre-computed operational R-D curve of each subband, the TCQ stepsize is determined by means of optimal bit allocation. Joshi, Fische, and Bamberger [4] apply the TCQ jointly with the optimal subband classification to wavelet image coding. Each subband is first classified into a few classes by a recursive optimization algorithm which maximizes the classification gain. Each class is then modeled as a generalized Gaussian source and encoded by the TCQ and arithmetic coding. These two algorithms are both based on the pre-computed R-D curves and optimal bit allocation which involves fairly high computational complexity.

In this work, we investigate a more efficient way to apply the TCQ to wavelet image coding with regard to both the computational complexity and the compression performance. This paper is organized as follows. In Section 2, we provide a brief review of the TCQ. Then, in Section 3, we propose a coding algorithm in which the subband data is first converted with blockwise binary classification to the type of source for which the TCQ has the best performance. Despite the greatly reduced computational complexity, the extensive simulation results presented in Section 5 show that the performance of the proposed algorithm is quite competitive when compared with other coding results reported in the literature.

2. TRELLIS CODED QUANTIZATION

The TCQ has the advantage of granular gain over the scalar quantization [1]. It can be regarded as a special vector quantization with structured codebook which is automatically generated from an expanded set of scalar quantization levels during the quantization process [?].

Let Δ be the stepsize of TCQ. The uniform output levels are then given by $\{P_i = i\triangle\}$. They are partitioned into four subsets $D_0 = \{P_{4i}\}$, $D_1 = \{P_{4i+1}\}$, $D_2 = \{P_{4i+2}\}$, and $D_3 = \{P_{4i+3}\}$ as shown in Fig. 1(a). An 8-state (S_1 to S_8) trellis as shown in Fig. 1(b) is defined by a rate-$\frac{1}{2}$ convolutional coder. There are two branches leaving each state. The top branch and the bottom one are marked by '0' and '1', respectively. Each branch is associated with one of the four subsets as shown in Fig. 1(b). For example, the top and the bottom branches leaving S_1 are associated with D_0 and D_2, respectively.

In the TCQ, to quantize an input sequence is to pick up a sequence of connected branches (path) in the trellis. To quantize a coefficient select one of the two branches which leave the current state. The state at the other end of the selected branch becomes the current state in turn. A branch is selected if the transform coefficient is quantized to one output level in the subset which is associated with itself. For example, at stage i, as shown in Fig. 1(b), the branch leaving A associated with D_1 is selected if the input data X_i is quantized to one output level, denoted with $Q(X_i)$, in the subset D_1. Certainly, to minimize the distortion, $Q(X_i)$ should be the closest output level to X_i in D_1. It should be noted that the quantization of X_i depends on the quantization of all the previous coefficients $\{X_j | 1 \leq j \leq i-1\}$. For a given sequence of data to be quantized, the TCQ encoder uses the Viterbi algorithm [5] to pick up the path that minimizes the mean square error (MSE) between the input data sequence $\{X_i\}$ and the output level sequence $\{Q(X_i)\}$.

Note that, in Fig. 1(b), the two branches leaving any trellis state are associated either with D_0/D_2 or with D_1/D_3, which implies that the quantization output of each transform coefficient is either from the super set $Z_0 = D_0 \cup D_2$ or $Z_1 = D_1 \cup D_3$. Therefore, the TCQ encoder only needs to send out the index of $Q(X_i)$ within the super set Z_0 or Z_1. For example, suppose the current position of the path is B whose state is S_3, the super set from which $Q(X_{i+1})$ can be chosen is $Z_0 = D_0 \cup D_2$. If the index of $Q(X_i)$ inside Z_0 is sent out, the decoder knows from which subset, either D_0 or D_2, the quantization level $Q(X_i)$ has been chosen from. In addition, the next state that can be decided which is either S_6 or S_5.

3. BLOCKWISE BINARY CLASSIFICATION AND DATA PARTITIONING

In order to apply the TCQ to wavelet image coding to achieve high compression efficiency, especially at very low bit rates, we use blockwise binary classification and data partitioning to convert the subbands data to the type of data for which the TCQ has the best quantization performance.

A. Blockwise Classification and Data Partitioning

After wavelet transform and before quantization, the image subbands are equally partitioned into 2×2 blocks. Each block is termed an *all-zero block* if the maximum magnitude of all the coefficients inside are less than a given threshold T. Otherwise it is termed a *non-zero block*. Each all-zero block or non-zero block is marked by a '0' or a '1', respectively. These block marks are compressed with the arithmetic coding and sent out to the decoder.

Each coefficient in the all-zero blocks is quantized to zero. Only the non-zero blocks are encoded by a 4-state uniform TCQ [1]. In the arithmetic coded TCQ (ACTCQ) [3], the quantization output is further compressed by an arithmetic coder. We observe that the compression performance of the arithmetic coder can be improved by adding variable length integer (VLI) coding before it. Suppose q_i is an output index of the TCQ. Its size S_i is defined as

$$S_i = \begin{cases} 0 & if \quad q_i = 0; \\ \lfloor \log_2 |q_i| \rfloor + 1 & if \quad q_i \neq 0. \end{cases} \tag{1}$$

If $q_i \neq 0$, its *residue bits* consists $S_i - 1$ bits for the binary representation of $|q_i| - 2^{S-1}$ and one additional bit for the sign. The size sequence $\{S_i\}$ is further compressed by a first-order arithmetic coder while the residue bits are sent out directly for the sake of low complexity The VLI coding improves the compression performance of the arithmetic coder by reducing the total bin number and enhancing the local correlation of the input data [7].

B. Determine the TCQ Stepsize

Next we determine the values of $\triangle$ and T. Obviously, their different vaules correspond to different bit allocation between the source of all-zero blocks and the source of all non-zero blocks. To maintain the low computational complexity of the proposed coding algorithm, we try to avoid introducing the optimal bit allocation into our coding algorithm when choosing the values of $\triangle$ and T. Let $\mathcal{R} = \frac{\triangle}{T}$. Table 1 shows the peak signal-to-noise ratio (PSNR) results for images Lena, Barbara and Goldhill at 0.25 bpp with different $\mathcal{R}$. It can be seen that $\mathcal{R} = 0.6$ always yields the near optimum compression performance. This can be explained as follows. The input of the TCQ is the 1-D immediate array formed by all the non-zero blocks. The variation of the statistical properties and spatial distribution of the 1-D array is largely reduced by the blockwise binary classification. This is because all the all-zero blocks which are essential to the variation of the spatial distribution and compression ratio are already "kicked out".

4. ADVANTAGES OF THE PROPOSED ALGORITHM

Fig. 2 shows the result of the blockwise partitioning and binary classification for Lena at $T = 25.4$. The white and the black regions represent the non-zero blocks and the all-zero blocks, respectively. The percentage of the non-zero blocks is 7.1%. The percentages of non-zero blocks of images Lena, Barbara, Goldhill, and Peppers at bit rates of 0.25 bpp and 0.50 bpp are listed in Table 2. In the proposed algorithm, the TCQ is only applied to the non-zero blocks instead of the whole decomposed image. Therefore, the computational complexity of the quantization has been greatly reduced by about 90%.

5. EXPERIMENTAL RESULTS

We have applied the proposed coding method to many images across a wide range of bit rates. Compression results of SPIHT [6], EZW [8], ACTCQ [3], and the proposed coding algorithm for 512×512 gray images Lena and Barbara are plotted in Figs. 3 and 4. The 9/7 wavelet filter and the 5-level dyadic subband decomposition are employed in the simulation. The experimental results show that the proposed method consistently outperforms the SPIHT by about 0.3 dB for Lena and 0.5 dB for Barbara.

Table 1: PSNR results for image Lena, Barbara and Goldhill at 0.25 bpp with different ratio $\mathcal{R}$.

$\mathcal{R} = \frac{\Delta}{T}$	Lena	Barb	Goldhill
0.3	33.97	27.45	30.19
0.4	34.21	27.73	30.40
0.5	34.33	27.93	30.49
0.6	34.37	28.03	30.55
0.7	34.33	28.04	30.51
0.8	34.29	28.01	30.44
0.9	34.19	27.94	30.33

Table 2: Percentage of blocks to be encoded at bit rates of 0.25 bpp and 0.50 bpp.

Bpp	Lena	Barb	Goldhill	Peppers
0.25	7.2%	8.8%	9.1%	6.9%
0.50	13.2%	14.3%	17.9%	14.8%

6. CONCLUSION

In this paper, we first provide a brief review of the TCQ and some TCQ-based wavelet image coding systems. In order to apply the TCQ to wavelet image coding to achieve better compression performance, especially at very low bit rate, we propose a new coding algorithm in which the TCQ is applied to non-zero blocks after blockwise binary classification. The proposed algorithm does not involves the pre-computation of the operational R-D curves and the optimal bit allocation. The computational complexity of the quantization has been reduced by more than 90%. However, the experimental results show that the compression performance of the proposed coding algorithm is quite competitive when compared with other results reported in the literature.

Acknowledgments

This work was supported in part by a University of California MICRO grant with matching support from Lucent Technologies, National Semiconductor, Tektronix Corporation, and Xerox Corporation.

7. REFERENCES

[1] M. W. Marcellin and T. R. Fischer, "Trellis coded quantization of memoryless and Gaussian-Markov sources," *IEEE Trans. on Communications*, vol. 38, pp. 82-93, January 1990.

[2] M. W. Marcellin, "On entropy-constrained trellis coded quantization," *IEEE Trans. on Communications*, vol. 42, pp. 14-16, January 1994.

[3] R. L. Joshi, V. J. Crump and T. R. Fischer "Image subband coding using arithmetic coded trellis coded quantization," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 5, pp. 515-523, December 1995.

[4] R. L. Joshi, T. R. Fischer, and R. H. Bamberger, "Optimum classification in subbabd coding of images," *Proc. International Conference Image Processing*, vol. 2, pp. 883-887, November 1994.

[5] G. D. Forney, Jr., "The Viterbi algorithm," *Proceding of IEEE*, vol. 61, pp. 268-278, March 1973.

[6] 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 System for Video Technology*, vol. 6, pp. 243-250, June 1996.

[7] Z. He, T. H. Yu and S. K. Mitra, "Block-wise zero mapping image coding," *International Conference on Image Processing*, Vancover, September 2000.

[8] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Proceesing*, vol. 41, pp. 3445-3462, December 1993.

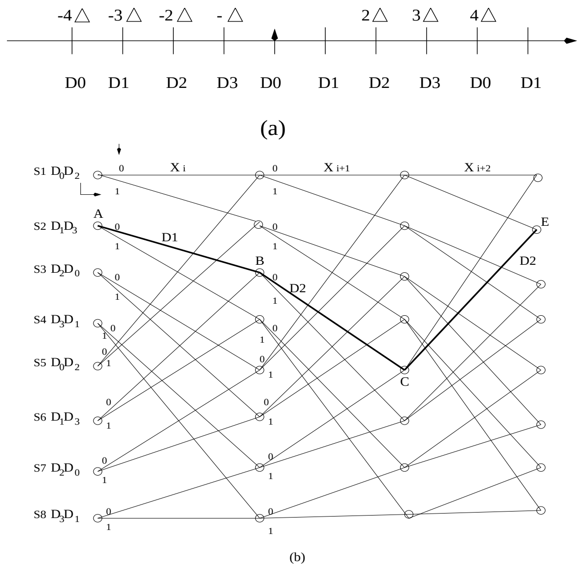

Figure 1: (a) Partition of the uniform codebook for the TCQ; (b) 8-state trellis for the TCQ.

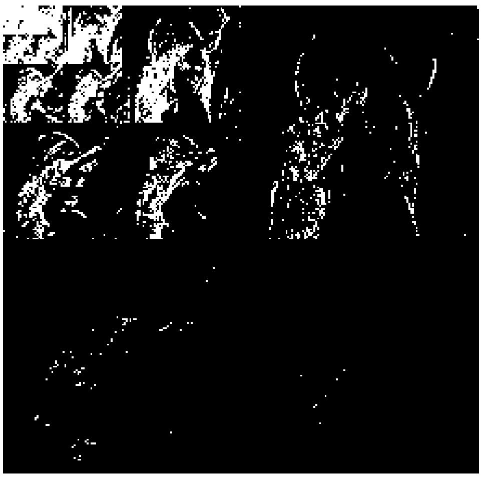

Figure 2: Blockwise partitioning and binary classification for Lena at0.25 bpp.

Table 3: PSNR results for image Lena, Barbara and Goldhill at 0.25 bpp with different ratio $\mathcal{R}$.

$\mathcal{R} = \frac{\Delta}{T}$	Lena	Barb	Goldhill
0.3	33.97	27.45	30.19
0.4	34.21	27.73	30.40
0.5	34.33	27.93	30.49
0.6	34.37	28.03	30.55
0.7	34.33	28.04	30.51
0.8	34.29	28.01	30.44
0.9	34.19	27.94	30.33

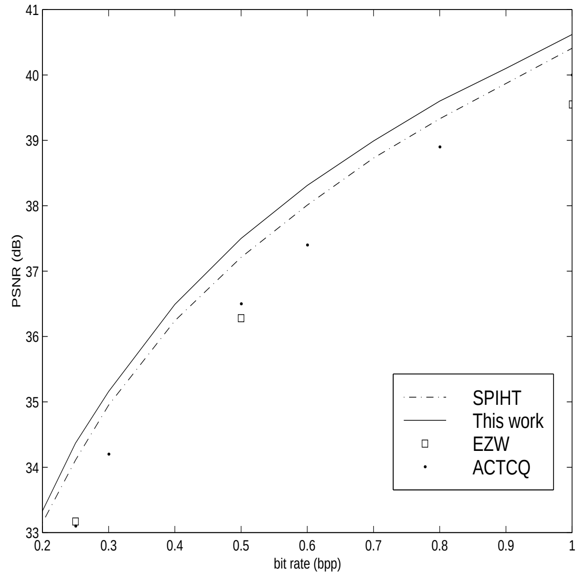

Figure 3: Rate-PSNR performance comparison of coding algorithms for image Lena.

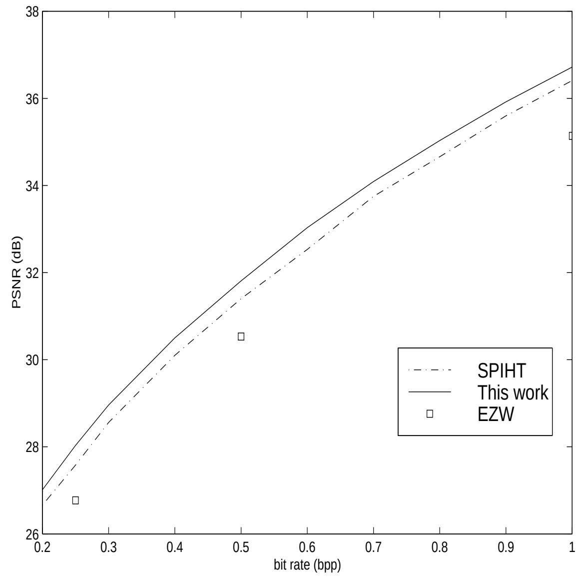

Figure 4: Rate-PSNR performance comparison of coding algorithms for image Barbara.