WAVELET-BASED IMAGE CODER WITH CHANNEL-OPTIMIZED TRELLIS-CODED QUANTIZATION

Tuyet-Trang Lam
Telecom. Research Center
Dept. of Electrical Engineering
Arizona State University
Tempe, AZ 85287-7206
lam@asu.edu

Glen P. Abousleman
Speech and Signal Processing Lab.
Motorola, SSG
8201 E. McDowell Rd, M/D H1176
Scottsdale, AZ 85257
glen1@kuma.geg.mot.com

Lina J. Karam
Telecom. Research Center
Dept. of Electrical Engineering
Arizona State University
Tempe, AZ 85287-7206
karam@asu.edu

ABSTRACT

This paper presents a wavelet-based image coder optimized for transmission over binary symmetric channels (BSC). The proposed coder uses a channel-optimized trellis-coded quantization (COTCQ) stage that is designed to optimize the image coding based on the channel characteristics. This optimization is performed only at the level of the source encoder, and does not include any channel coding for error protection. Consequently, the proposed channel-optimized image coder is especially suitable for wireless transmission due to its reduced complexity. Furthermore, the improvement over TCQ-based image coders is significant. Examples are presented to illustrate the performance of the proposed COTCQ-based image coder.

1 INTRODUCTION

Numerous techniques have been developed for coding still images and video. However, most of the existing image coders have been developed under the assumption of reliable noise-free transport. Therefore, the transmitted data may be corrupted by channel noise, and the quality of the reconstructed data may be degraded substantially unless redundant information or error-correcting methods are added.

Trellis-coded quantization (TCQ) has been used in several coding systems designed to operate in noisy channel environments (e.g., [1]). In fact, TCQ has been shown [2] to provide greatly improved rate-distortion performance as compared to scalar quantization. It also provides significant robustness to channel errors as compared to vector quantization (VQ). In addition, TCQ is much less computationally complex than VQ.

In this paper, we propose a source coder that performs well in a noisy environment without any channel coding or error correcting schemes. The proposed image coder uses a 22-band wavelet decomposition, where each sub-band is quantized using a channel-optimized trellis-coded quantizer (COTCQ).

This paper is organized as follows. Section 2 provides a brief overview of TCQ. The proposed channel-optimized image coder is described in Section 3. Coding examples are presented in Section 4. A conclusion is given in Section 5.

2 TRELLIS-CODED QUANTIZATION

Trellis-coded quantization (TCQ) was developed in [2]. As mentioned in Section 1, TCQ has several performance advantages over scalar and vector quantization. It also has the nice property of recovering from bit errors [2]. For example, one bit cannot affect more than $1 + \log_2(N)$ outputs, where N is the number of trellis
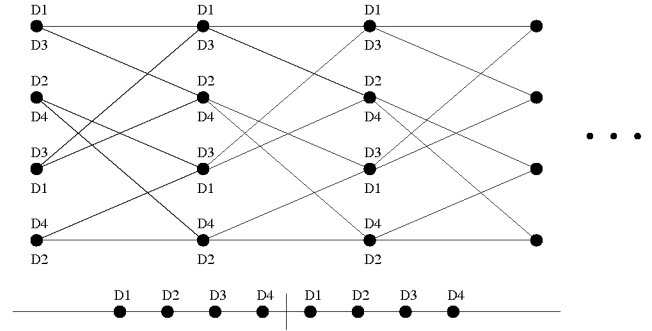

Figure 1. A four-state trellis with subset labeling.

states. This is one advantage of TCQ as compared to other coding methods, such as VQ-based methods, where a single channel error can destroy a block of data.

For encoding a memoryless source at R bits per sample using TCQ, a codebook of size 2^{R+1} is partitioned into four subsets D_j ($j = 0, 1, 2, 3$), each containing 2^{R-1} codewords. These subsets are used as labels for the branches of a suitably chosen trellis. An example is shown as shown in Fig. 1 for $R = 2$. Given an input data sequence $\mathbf{x} = \{x_1, x_2, \ldots\}$, the best (in the mean-squared-error sense) allowable sequence of codewords is determined as follows. For the ith stage in the trellis (corresponding to x_i), the best codeword in the jth subset ($j = 0, 1, 2, 3$), say c_j, is chosen, and the associated cost $\rho_j = (x_i - c_j)^2$ is calculated. Each branch in the ith stage of the trellis that is labeled with subset D_j is assigned cost ρ_j. The Viterbi algorithm [3] is then used to find the trellis path with the lowest overall cost.

3 CHANNEL-OPTIMIZED TCQ-BASED WAVELET IMAGE CODER

A block diagram of the proposed channel-optimized TCQ (COTCQ) based wavelet image coder is shown in Fig. 2. The input image is first decomposed using a 9-7 biorthogonal 2-D Discrete Wavelet Transform (DWT) into 22 sub-bands as in [4]. The statistics of each sub-band are computed. All sub-bands are normalized by subtracting their mean (only the sequence corresponding to the lowest-frequency sub-band is assumed to have a non-zero mean) and dividing by their respective standard deviations. The normalized sub-bands are then encoded using a fixed-rate COTCQ system designed for the memoryless (zero mean and unit-variance) Laplacian source. In addition, an optimal (in the MSE sense) rate allocation scheme is utilized.
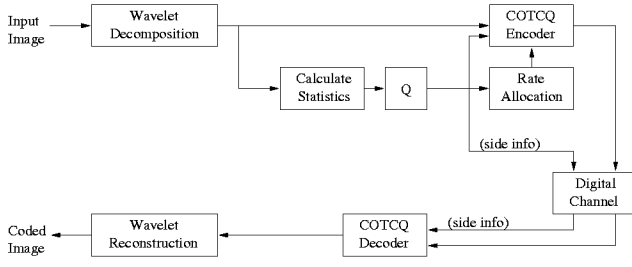
Figure 2. Block diagram of the proposed wavelet image coder using COTCQ.

The encoder transmits the mean of the lowest-frequency subband and the standard deviations of all 22 sub-bands as side information. These quantities are quantized with 16-bit uniform scalar quantizers using a total of 368 bits. The initial trellis state for each encoded sub-band and the target compression ratio must also be transmitted and are allotted 2 bits per sub-band (for a 4-state trellis) and 10 bits, respectively. Therefore, the total side information consists of 422 bits per image. The decoder performs the inverse operations and consists of a COTCQ decoder followed by a wavelet synthesis stage.

The wavelet decomposition, COTCQ, and rate allocation stages are discussed below.

3.1 Wavelet Decomposition

The input image is transformed using a 9-7 biorthogonal 2-D DWT [5] into 22 sub-bands in a modified Mallat tree configuration. That is, the image is initially decomposed into 16 equal-sized sub-bands, with two additional levels of decomposition being applied to the lowest-frequency sub-band.

3.2 Channel-Optimized Trellis-Coded Quantization (COTCQ)

We designed a channel-optimized trellis-coded quantization system for use with the binary symmetric channel (BSC) with a bit error probability P_b. For the considered image coding application, we extended the work in [6] by deriving a general expression of the transition probability matrix (Section 3.2.1) in terms of the encoding bit rate R (only $R = 1$ and $R = 2$ are considered in [6]). The design was performed to support the following bit error probabilities:

$$P_b = 0.0,\ 0.0001,\ 0.0005,\ 0.001,\ 0.005,\ 0.01,\ 0.05,$$
$$0.1,\ 0.15,\ 0.2,\ 0.25,\ 0.3,\ 0.35,\ 0.4,\ 0.45,\ 0.5 \quad (1)$$

For each P_b, fixed-rate COTCQ codebooks were designed in one-bit increments from $R = 1$ to 8 bits per sample. For each sub-band i ($i = 1\ldots, 22$), the rate allocation scheme (Section 3.3) uses the computed sub-band statistics and selects a R_i-bit COTCQ codebook, $0 \leq R_i \leq 8$, from the designed COTCQ codebooks, so that the overall distortion is minimized while maintaining the specified overall bit rate R_T.

In what follows, let R denote the encoding bit rate (corresponding to one COTCQ quantizer), P_b the bit error probability of the transmission channel, and N_1 the number of TCQ subsets. In our case, since we consider a 4-state trellis, $N_1 = 4$. The COTCQ codebook used to encode a sequence at R bits/sample contains 2^{R+1} codewords. Each of the 4 subsets then contains $N_2 = 2^{R-1}$ codewords.

Consider an input sequence $X = \{x_j\}_{j=1}^{\|x\|}$, where $\|x\|$ denotes the number of elements in the sequence X. Let $Y = \{y_{k,l}\}$ be the set of reconstructed levels, or codewords. $y_{k,l}$, $0 \leq k \leq 4$, $0 \leq l \leq 2^{R-1}$, denotes the codeword corresponding to the lth level of the kth subset D_k. We define the distortion measure between x_j and $y_{k,l}$ as:

$$d(x_j, y_{k,l}) = \sum_{i=0}^{N_1-1} \sum_{n=0}^{N_2-1} P_{i,n/k,l}(x_j - y_{i,n})^2 \quad (2)$$

where $P_{i,n/k,l} = Pr\{y_{i,n}$ received / $y_{k,l}$ is sent $\}$ and is given by the transition probability matrix derived in Section 3.2.1. We define the overall distortion as:

$$D = \frac{1}{\|x\|} \sum_{x_j \in X} d(x_j, y_{k,l}) \quad (3)$$

where x_j has been encoded as $y_{k,l}$, after the COTCQ has found its path through the trellis to minimize the overall distortion D.

For a given bit error probability P_b and an encoding rate R, the quantization codebook is designed using the K-means algorithm [7] and the defined distortion metric (3). At each iteration, a training sequence is encoded using the metric described above and clustered. The cluster centroids are then used to encode the training sequence and to update the old centroids until the algorithm converges to a local optimum.

Let m be an iteration index. Let $Y^{(m)} = \{y_{k,l}^{(m)}\}$ be the set of reconstruction levels at the mth iteration, $D^{(m)}$ be the overall distortion at the mth iteration, and $Q_{k,l}^{(m)} = \{x_j : x_j \in X$ is encoded as $y_{k,l}^{(m)}\}$. The algorithm used to design the optimum codebooks is described as follows [6]:

- Step 0 - Initialization: set $m = 0$ and the overall distortion measure $D^{(-1)} = \infty$. The training sequence, the convergence threshold ϵ, and the initial reproduction codebook are given.

- Step 1 - Reproduction Codebook Update: encode the training sequence, using the Viterbi algorithm and the codebook $\{y_{k,l}^{(m)}\}$, to obtain $\{Q_{k,l}^{(m)}\}$ and $D^{(m)}$ as given by (3). If $(D^{(m-1)} - D^{(m)})/D^{(m)} < \epsilon$, quit; otherwise, update the centroids as follows:

$$y_{i,n}^{(m+1)} = \frac{\sum_{x_j \in Q_{k,l}^{(m)}} x_j}{\sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} P_{i,n/k,l}\|Q_{k,l}\|^{(m)}} \quad (4)$$

- Step 2 : set $m = m + 1$, and go to Step 1.

For our image coder, we design a set of channel-optimized trellis-coded quantizers for values of R ranging from 1 to 8 bits per sample, and the values of P_b as indicated in (1). For $P_b = 0.0$ (which corresponds to regular TCQ since no noise is present), the initial codebook in Step 0 is chosen to contain the reconstruction levels of a scalar quantizer. For the subsequent P_b values, we follow the method described in [8], which consists of using, as the initial codebook, the optimum codebook that was found for the preceding P_b value.
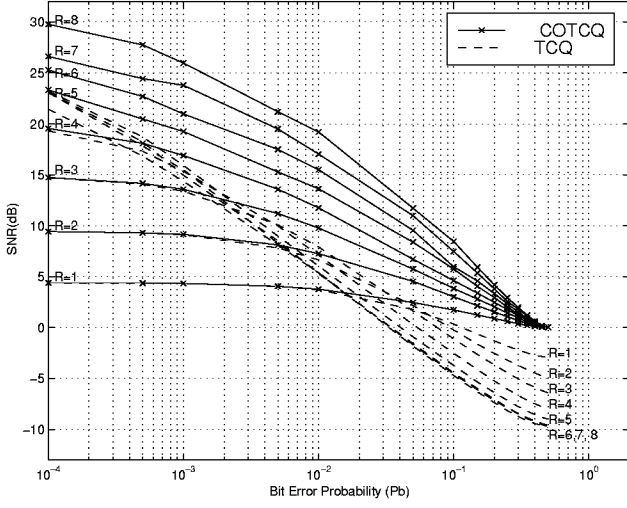
Figure 3. Performance of COTCQ vs. TCQ for Laplacian sources at encoding rates R = 1 to 8 bits/sample.

3.2.1 Transition probability matrix

The entries of the transition probability matrix consist of the transition probabilities $P_{i,n/k,l} = \text{Prob}\{y_{i,n} \text{ received} / y_{k,l} \text{ sent}\}$, where $y_{k,l}$ is generated by the encoder and sent through the binary symmetric channel. i, k correspond to the selected subsets, and n, l to the codewords chosen from subsets i and k, respectively. The transition probability matrix is characterized by the channel bit error probability P_b.

The transition probability matrix for $R = 1$ is [6]:

$$\begin{bmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{bmatrix} \quad (5)$$

where :

$$\begin{aligned} a &= P_{0,0/0,0} = (1 - P_b)^3 P_b + (1 - P_b)P_b^2 \\ b &= P_{1,0/0,0} = (1 - P_b)^2 P_b + P_b^3 \\ c &= P_{2,0/0,0} = 2(1 - P_b)^2 P_b \\ d &= P_{3,0/0,0} = 2(1 - P_b)P_b^2 \end{aligned} \quad (6)$$

This matrix is symmetric with respect to its two axes, due to the structure of the trellis.

Let $b_{(i,n)}$ and $b_{(k,l)}$ denote the $(R - 1)$-bit binary code of codeword n of subset i, and codeword l of subset k, respectively. For higher bit rates, it can be shown that the transition probability $P_{i,n/k,l}$ is proportional to $P_{i,0/k,0}^{(R=1)}$, and that, for any bit rate R, it is given by:

$$P_{i,n/k,l} = P_{i,0/k,0}^{(R=1)}(1 - P_b)^{\#same} P_b^{\#different} \quad (7)$$

$\#different$ and $\#same$ in (7) are given by:

$$\begin{cases} \#different &= \|b_{(i,n)} \ xor \ b_{(k,l)}\|_{\text{Hamming}} \\ \#same &= R - 1 - \#different. \end{cases} \quad (8)$$
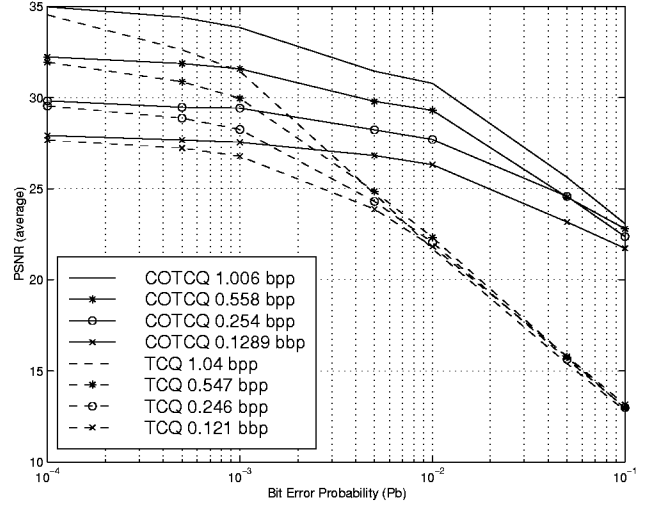

Figure 4. Performance of the proposed COTCQ-based image coder compared to the TCQ-based image coder.

where $\|.\|_{\text{Hamming}}$ is the Hamming distance.

For example, with a, b, c and d as defined in (5), (7) gives the following transition probability matrix for $R = 2$:

$$\begin{array}{c} \\ P_{0,0} \\ P_{0,1} \\ P_{1,0} \\ P_{1,1} \\ P_{2,0} \\ P_{2,1} \\ P_{3,0} \\ P_{3,1} \end{array} \begin{array}{cccccccc} P_{0,0} & P_{0,1} & P_{1,0} & P_{1,1} & P_{2,0} & P_{2,1} & P_{3,0} & P_{3,1} \\ \begin{bmatrix} a\overline{P_b} & aP_b & b\overline{P_b} & bP_b & c\overline{P_b} & cP_b & d\overline{P_b} & dP_b \\ aP_b & a\overline{P_b} & bP_b & b\overline{P_b} & cP_b & c\overline{P_b} & dP_b & d\overline{P_b} \\ b\overline{P_b} & bP_b & a\overline{P_b} & aP_b & d\overline{P_b} & dP_b & c\overline{P_b} & cP_b \\ bP_b & b\overline{P_b} & aP_b & a\overline{P_b} & dP_b & d\overline{P_b} & cP_b & c\overline{P_b} \\ c\overline{P_b} & cP_b & d\overline{P_b} & dP_b & a\overline{P_b} & aP_b & b\overline{P_b} & bP_b \\ cP_b & c\overline{P_b} & dP_b & d\overline{P_b} & aP_b & a\overline{P_b} & bP_b & b\overline{P_b} \\ d\overline{P_b} & dP_b & c\overline{P_b} & cP_b & b\overline{P_b} & bP_b & a\overline{P_b} & aP_b \\ dP_b & d\overline{P_b} & cP_b & c\overline{P_b} & bP_b & b\overline{P_b} & aP_b & a\overline{P_b} \end{bmatrix} \end{array} \quad (9)$$

where $\overline{P_b} = 1 - P_b$. From (9), it can be seen that each 2×2 block follows the pattern given by the transition matrix for $R = 1$. Using (7), the transition probability matrix can be derived in a similar way for other values of R.

3.2.2 Simulation results for COTCQ

The performance of COTCQ relative to TCQ is illustrated in Fig. 3 for Laplacian sources. Fig. 3 shows the Signal-to-Noise Ratio (SNR) for COTCQ (solid lines) and TCQ (dashed lines) as a function of P_b, for $R = 1$ (bottom solid and dashed plots) to 8 (top solid and dashed plots), in one-bit increments. A training sequence of 100,000 samples was used for designing the COTCQ and TCQ codebooks. The plots of Fig. 3 are generated by encoding a Laplacian sequence of 100,000 samples (different from the training sequence) using the designed quantization codebooks.

At $P_b = 0$ (not shown in Fig. 3), COTCQ and TCQ have identical performance (the COTCQ and TCQ curves intersect at $P_b = 0$). However, as P_b increases, the performance of TCQ degrades at a significantly faster rate as compared to COTCQ. Also, as R increases, the relative degradation rate of TCQ increases, and the TCQ performance curves diverge significantly from the COTCQ curves even at values of P_b as low as 10^{-4}. Note that, for $R = 1$ to 3, the TCQ and COTCQ performance curves are almost iden-

(a) TCQ wavelet image coder
PSNR = 21.93 dB

(b) Proposed COTCQ wavelet image coder
PSNR = 29.13

Figure 5. 512×512 Lena image encoded at 0.5 bpp with a channel bit error probability $P_b = 0.01$.

tical for low values of P_b, but diverge as P_b increases, with TCQ degrading at a faster rate than COTCQ.

3.3 Rate Allocation

Rate allocation is performed by using the iterative algorithm described in [9]. It allocates bits based on the energy content of the wavelet coefficients, and on the rate-distortion performance of the designed channel-optimized trellis-coded quantizers (solid curves in Fig. 3). This rate allocation procedure allows precise bit rate specification, independent of the image to be coded.

4 CODING EXAMPLES

Fig. 4 illustrates the performance of the proposed COTCQ-based wavelet image coder (solid lines) for different encoding rates, $R_T = 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$, and different bit error probabilities P_b. For comparison, Fig. 4 also shows the performance of the same wavelet image coder, but with the COTCQ stage replaced by a TCQ stage (dashed lines). The performance plots of Fig. 4 are obtained by averaging, over ten experiments, the Peak SNR curves corresponding to coding an input image. Thus, for each R_T and P_b, ten experiments are conducted, where each experiment consists of encoding the image, corrupting the resulting bit stream with noise at probability P_b, then decoding the image and computing the resulting PSNR.

Figs. 5(a) & (b) show the 512 × 512 Lena image encoded at $R_T = 0.5$ bit per pixel (bpp), with a bit error probability $P_b = 0.01$, using the TCQ-based image coder and the proposed COTCQ-based image coder, respectively. The Peak SNR is 29.13 dB using the COTCQ-based image coder (Fig. 5(b)) and decreases to 21.93 dB if TCQ is used in place of COTCQ (Fig. 5(b)). Moreover, the TCQ-based image coder produces annoying noisy spots in the reconstructed image(Fig. 5(a)). Most of these spots are not present when the COTCQ-based image coder is used (Fig. 5(b)). Also, compared to other TCQ-based image coders designed to operate in noisy channel environments (e.g., [1]), the proposed image coder exhibits improved coding performance despite its low complexity.

5 CONCLUSION

The proposed image coder facilitates the transmission of high-quality compressed imagery over noisy channels without the need for channel coding. In addition, no entropy coding is used, which further reduces the complexity and susceptibility to channel errors. The simplicity and good performance of the proposed image coder in the presence of noise make the presented coder especially suitable for wireless communication applications.

REFERENCES

[1] Z. Sun, C. W. Chen, and K.J. Parker, "Image transform coding using trellis coded quantization through noisy channels", *Proceedings of the SPIE*, vol. 2915, pp. 13–24, 1997.

[2] M. W. Marcellin and T. R. Fischer, "Trellis coded quantization of memoriless and gauss-markov sources", *IEEE Trans. Communication*, vol. 38, pp. 82–93, 1990.

[3] Jr. G. D. Forney, "The viterbi algorithm", *Proc. IEEE*, vol. 61, pp. 268–278, Mar. 1973.

[4] G. P. Abousleman, "Wavelet-based hyperspectral image coding using robust fixed-rate trellis coded quantization", *Proceedings of the SPIE*, vol. 3372, pp. 74–85, 1998.

[5] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transforms", *IEEE Trans. on Image Processing*, vol. 1, pp. 205–220, Apr. 1992.

[6] M.Wang and T. R.Fischer, "Trellis coded quantizaion designed for noisy channels", *IEEE Trans. Information Theory*, vol. 40, pp. 1792–1801, 1994.

[7] J.T Tou and R.C. Gonzalez, *Pattern Recognition Principles*, vol. 1, chapter 3, 1974, Pattern Classification by Distance Functions.

[8] E. Ayanoglu and R. M. Gray, "The design of joint source and channel trellis waveform coders", *IEEE Trans. Information Theory*, vol. 33, pp. 855–865, 1987.

[9] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers", *IEEE Trans. Acoust., Speech, and Signal Proc.*, vol. 36, pp. 1445–1453, Sept. 1988.