JOINT SOURCE CHANNEL CODING OF IMAGES WITH TRELLIS CODED QUANTIZATION AND CONVOLUTIONAL CODES

Benjamin Belzer[1], John D. Villasenor[1], and Bernd Girod[2]

[1]UCLA Electrical Engineering Department
405 Hilgard Avenue
Los Angeles, CA 90024-1594

[2]Lehrstuhl für Nachrichtentechnik
University of Erlangen-Nuremberg, Germany

ABSTRACT

The design of a low complexity joint source channel codec for images using trellis coded quantization and convolutional codes is studied. It is shown that a low complexity joint source channel codec offering high robustness is enabled by using the same Ungerboeck trellis code to do both quantization and convolutional coding. Simulation results are presented for joint source channel coding of Gaussian sources over the additive white Gaussian noise channel (AWGN) under the restriction of BPSK modulation; coding of image subband coefficients is also performed. The results compare favorably with those previously reported for systems employing unified trellis source coding and convolutional channel coding, and for pseudo-Gray coded systems. The simulation results also show that there is an optimal allocation of transmission bandwidth between the source and channel codes.

1. INTRODUCTION AND BACKGROUND

Mobile multimedia personal communications systems require joint source channel coding (JSCC) in order to maximize quality of service under rapidly varying channel conditions. JSCC codecs must offer a large array of source rates to ensure optimal encoding of sources with varying information content, and they must offer a wide range of available error protection levels in order to transmit information of varying error sensitivity over rapidly changing radio channels. At the same time, low power portable systems require low complexity solutions. A critical part of JSCC systems is the mapping between quantized source symbols and their representative channel symbols. A particularly elegant approach that offers both reasonable complexity and (in the case of Gaussian channels) optimal source/channel mapping is the joint trellis coded quantization/modulation (TCQ/TCM) method of [1], in which the same Ungerboeck trellis [3] is used for the TCQ and TCM. Because TCM with a wide choice of channel coding rates might be prohibitively complex for low power systems, it is attractive to also consider simpler systems that replace TCM with the combination of digital convolutional coding and BPSK. In this paper, the same Ungerboeck trellis is used for TCQ source coding (as in [2]) and for soft-decision digital convolutional channel coding. In this context, Section 2 below presents (a) methods to ensure the best possible mapping between Euclidean distance distortion in the source codec and Hamming distance in the channel codec, and (b) the performance improvement enabled by training the output levels by using the properties of the source and channel as opposed to those of the source alone. We present in Section 3 experimental results obtained when this technique is applied to coding of Gaussian sources as well as to coding of subband coefficients of the 512 x 512 Lena image.

2. JOINT OPTIMIZATION OF THE TRELLIS SOURCE AND CHANNEL CODES

By conditioning the mean squared error expectation on the transmitted and received quantizer outputs C_i and C_j, respectively, one can show that the total expected distortion D_{tot} between transmitted and received signals splits into source dependent and source/channel dependent terms D_s and D_{sc}, respectively, where

$$D_s = \sum_i p_i \int p(x|C_i)\,(x-C_i)^2 dx \tag{1}$$

and

$$D_{sc} = \sum_i p_i \sum_j p(j|i)\,(C_i - C_j)^2 + 2\sum_i p_i \sum_j p(j|i)\,(C_i - C_j)\int p(x|C_i)\,(x - C_i)\,dx \tag{2}$$

Here p_i refers to the probability that source symbol C_i is selected to represent source symbol x, and $p(j|i)$ is the probability of receiving C_j when C_i was actually sent. Note that the probabilities p_i depend on the trellis as well as on the source p.d.f. and the location of the C_i; thus the p_i cannot be calculated by a simple integration of the source p.d.f. over some quantization interval. The trellis processing also orces the inclusion of the second term in equation (2), since in general the C_i will not be the centroids of their quantization sets. Because the convolutional encoder used for channel encoding is also used for TCQ decoding, source symbol C_i is represented in the channel code by the binary expansion of its subset label i, so that the channel transition prob-

Figure 1: Subset labels and optimized output levels for 1 bit/sample symmetric codebook TCQ for a generalized Gaussian source with a shape parameter of 0.5, for 0 channel BER.

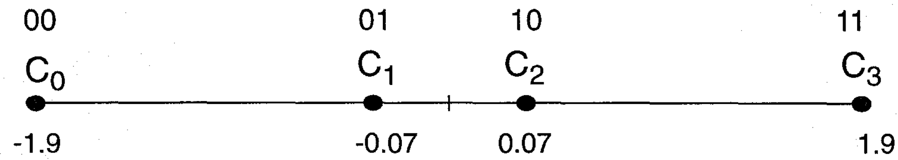
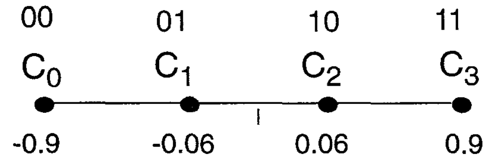


Figure 2: Joint source/channel optimized output levels for same source as Figure 1 transmitted over a Binary Symmetric Channel with BER = 10^{-1}; the binary representations of the subset labels are the transmitted channel symbols.



abilities $p(j|i)$ are determined by the Hamming distance between the subset labels. In this manner a mapping between source and channel symbols is achieved. For example, Figure 1 shows the TCQ 1 bit/sample output array, with the output levels optimized for the generalized Gaussian source with a shape parameterof $\nu = 0.5$. The maximum possible Hamming distance of 2 between C_0 and C_3 means that the probability $p(0|3)$ of confusing these two symbols in the channel is p^2, where p is the BSC crossover probability. Thus the two source symbols which would cause the most MSE distortion if confused with one another are assigned the minimum possible probability of error in the channel code. In the presence of a non-zero channel BER, minimization of the total distortion is achieved by training the output levels C_i (and with them their probabilities p_i) in a simulation which includes the end-to-end process of TCQ encoding, channel encoding, insertion of channel errors in the AWGN, soft decision channel decoding, and TCQ decoding. We use the training algorithm developed in [4] for joint TCQ/TCM systems. In this algorithm, which is basically a modified LBG scheme, the transmitted source words are clustered around their representative output words at the receiver, and the updated output words are the centroids of the clusters. Figure 2 shows that in the presence of a noisy channel the system optimizes its performance by reducing the energy (and potential MSE distortion) of the highest energy output words. We note that such a system assumes that the transmitter has some information about the channel SNR; such information is commonly available in practical communication systems.

2.1. TCVQ Design for JSCC

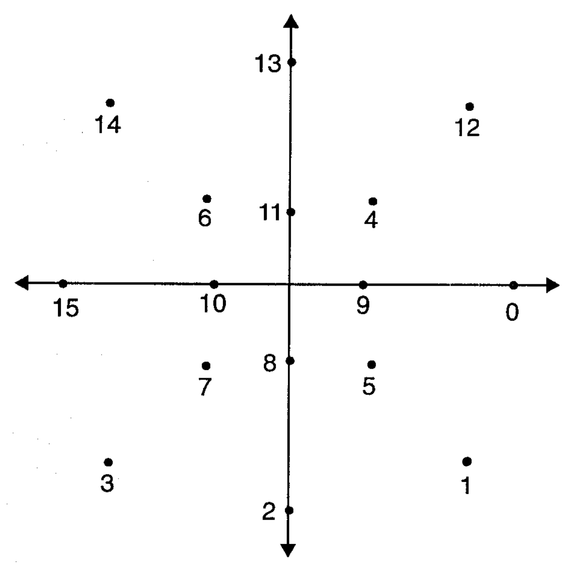
In order to source code at non-integer rates, and to handle highly peaked long tailed sources, it is necessary to code multiple samples per trellis branch; the resulting hybrid between VQ and trellis coding is called TCVQ. We follow the approach of [5] for partitioning the codebook into subsets and assigning the subsets to trellis branches. However, we optimize the subset labels so that the Hamming distance between output points is as proportional as possible to the Euclidean distance. This corresponds to optimization of the $p(j|i)$ in equation 2. Figure 3 shows our 2D TCVQ codebook for rate 1.5 bps coding of Gaussian i.i.d. sources. A 16 state rate 3/4 trellis is used; 3 bits are released for every 2 input samples. These three bits are then fed to the rate 3/4 convolutional encoder so that 4 bits per 2D input vector are sent over the channel; the four bits are the binary expansion of the subset labels shown in figure 3. When optimizing the subset labels, care must be taken not to destroy the distance properties of the channel codec. The dmin of our distance mapping optimized channel code is 3, only slightly smaller than the dmin of 4 for the best 16 state rate 3/4 codes.

We also designed a 1 bps coder for generalized Gaussian sources of shape parameter 0.5; such sources require higher dimension codebooks to realize the VQ shape gain. The codec uses a 4 dimensional symmetric codebook with 352 output points. Symmetrizing the codebook means that only 54 squaring operations need be calculated for each input sample; the use of 10 bit pre-quantization together with a lookup table enables muliplierless source coding. The coder releases 4 bits per input vector; the rate of the resulting channel code is 4/log2(352) = 0.4728.

3. SIMULATION RESULTS AND DISCUSSION

Figure 4a shows the SNR versus channel BER for jointly optimized output levels when a Gaussian i.i.d. source is encoded at 1bpp using a rate 1/2 32 state "Z_1" Ungerboeck code [3]. The 1bps source codebook for the Gaussian source appears much like those shown in figures 1 and 2, except that the optimal source levels are different (for a perfect channel they are +- 0.33 and +- 1.21). The channel is AWGN and the channel coder uses bit-wise soft decision decoding. The BER numbers on the x axis are calculated assuming that BPSK modulation is used. The figure also shows the curve for the case in which the output levels are optimized for the source coding only. As could be expected, the extent of the improvement enabled by joint optimization is greater for channels with higher BER. Also

Figure 3: Two dimensional symmetric codebook for 1.5 bps TCVQ of i.i.d. Gaussian sources. Point labels map Euclidean distance to Hamming distance.



shown on the graph are results reported by Freeman et al. [7] for the AWGN channel with Gaussian noise and soft decision decoding. The codec in [7] requires 12 multiplies/sample as opposed to the 2 multiplies/sample needed in the codec presented here. Figure 4b shows SNR vs. BER for the combination 1.5 bps TCVQ and rate 3/4 channel coder discussed in section 2.1. The benefits of joint source/channel training of the quantizer output points are more obvious at higher source bitrates and with less powerful channel codes.

Figure 5 considers the question of the best tradeoff between source and channel coding given a fixed channel transmission bandwidth per source sample. All of the curves in figure 5 have a transmission bandwidth of 2 bits per source sample. Curves o, * and x dedicate 2 bps, 1.5 bps, and 1 bps to the source code and the rest to the channel code. From the figure it is clear that the best strategy is to follow the convex hull of the three curves, switching source and channel rates when necessary. The distortion-rate function for the Gaussian source evaluated at the BSC channel capacity serves as an upper bound to the two source code only curves o and +. Curve + is the pseudo-Gray coded 2 bpp case as presented in [6], in which full-search VQ of vector dimension 6 and a 64 vector codebook with optimized index switching were used. The performance of pseudo-Gray coding can be improved at no complexity cost by using channel optimized VQ codebooks. However, further improvements require addition of a separate channel coder with its attendant hardware complexity, whereas the trellis based system uses the same hardware for both source and channel coding.

Figure 6 shows the SNR versus BER for subband coefficients from an octave band decomposition of the 512x512 Lena image; the 128x128 vertical edge detail signal was used as our data source. For training the quantizer output points, the subband coefficients were modelled with a generalized Gaussian source with a shape parameter of 0.5. The long tails of this distribution cause the optimal 0-BER codebook to have a much larger dynamic range than the codebooks designed for Gaussian sources. The high dynamic range leads to large gains when joint/source channel training of the codebook is performed; at higher BERs the outlying points are pulled in towards the origin by the optimization algorithm. For high BER values, the gain due to source/channel training ranges from 1.3 dB at 0.1 BER to 4.8 dB at 0.2 BER.

4. ACKNOWLEDGMENT

This work was supported by the U.S. Department of Justice/Federal Bureau of Investigation, ARPA/CSTO under Contract J-FBI-93-112, Computer Aided Design of High Performance Network Wireless Networked Systems, and by Hughes Aircraft Company and Texas Instruments. The authors would also like to acknowledge Benjamin Tang and Professor Greg Pottie, both of UCLA, for providing the Viterbi decoder used in this work and for useful discussions on trellis coding.

5. REFERENCES

[1] T.R. Fischer and M.W. Marcellin, "Joint Trellis Coded Quantization/Modulation", *IEEE Trans. on Communications*, Vol. 39, No. 2, February 1991, pp. 172-176.

[2] M.W. Marcellin and T.R. Fischer, "Trellis Coded Quantization of Memoryless and Gauss-Markov Sources", *IEEE. Trans. on Communications,* Vol. 38, No. 1, January 1990, pp. 82-93.

[3] G. Ungerboeck, "Trellis-Coded Modulation with Redundant Signal Sets Part II: State of the Art", *IEEE Communications Magazine,* Vol. 25, No. 2, February 1987, pp. 12-21.

[4] M. Wang and T.R. Fischer, "Trellis-Coded Quantization Designed for Noisy Channels", *IEEE Trans. on Information Theory,* Vol. 40, No. 6, November 1994, pp. 1792-1802.

[5] T.R. Fischer, M.W. Marcellin, and M. Wang, "Trellis-Coded Vector Quantization", *IEEE Trans. on Information Theory,* Vol. 37, No. 6, November 1991.

[6] K. Zeger and A. Gersho, "Pseudo-Gray Coding", *IEEE Trans. on Communications,* Vol. 38, No. 12, December 1990, pp. 2147-2158.

[7] G.H. Freeman, J.W. Mark, and I.F. Blake, "Trellis Source Codes Designed by Conjugate Gradient Optimization", *IEEE Trans. on Communications,* Vol. 36, No. 1, January 1988, pp. 1-12

Figure 4a: SNR vs. BER for coder using source/channel training (solid line) and source training only (dashed line). The transmitted bandwidth is two channel bits per source sample. The source coding bit rate is 1 bit/sample and the source is Gaussian. The rate 1/2 convolutional channel coder uses bitwise soft-decision decoding over the AWGN channel and transmits two channel bits for every source bit. Also shown are results from ref. [7], obtained using trellis source coding with soft decision channel decoding.

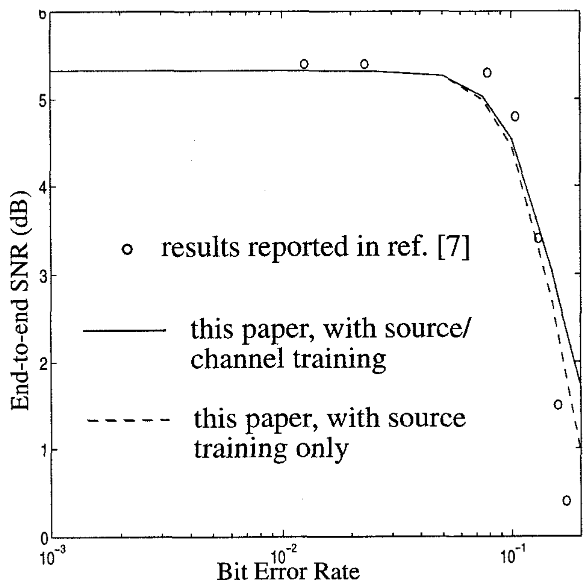


Figure 4b: A different tradeoff of bandwidth between source and channel coding: here as in figure 4a the total transmitted bandwidth is two channel bits per source sample, but the source coding rate is 1.5 bps and the channel coding rate is 3/4.

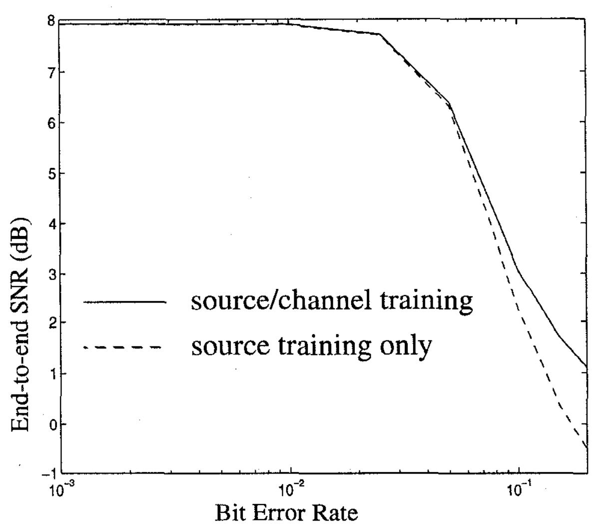


Figure 5: Summary of results for trellis based JSCC of Gaussian samples shows that there is an optimal tradeoff of transmission bandwidth between source and channel codes, depending on the BER of the AWGN channel. All curves shown have a transmission bandwidth of 2 bits per source sample. At low BER, dedicating both bits to the source coder is best (curve o). At higher BERs, expending first 1/2 and then 1 bps on channel coding is best (curves * and x).

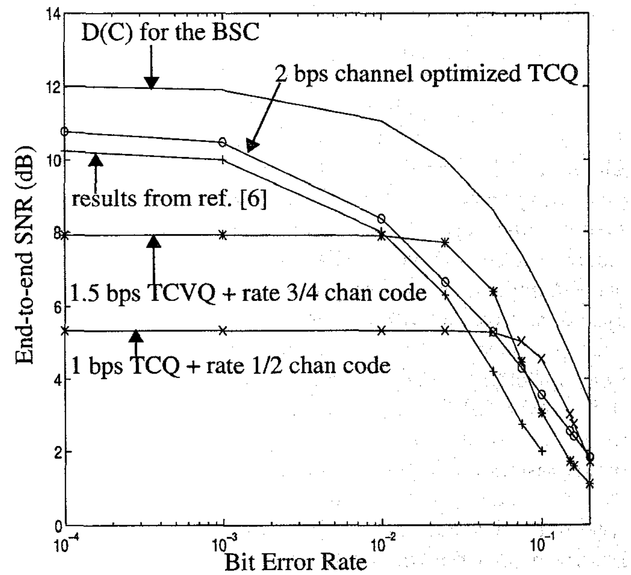


Figure 6: SNR vs. BER for subband coefficients from the level 2 vertical detail signal of the 512x512 Lena image. .