EUSIPCO-88

SUBBAND CODING OF IMAGES USING A TRELLIS CODER

Efrat Rotem and David Malah

Department of Electrical Engineering
Technion — Israel Institute of Technology
Haifa 32000, Israel

Abstract – In this paper, an image subband coding (SBC) system, which applies a trellis coder to efficiently encode the subband signals, is presented. The system uses a bank of two-dimensional quadrature mirror filters (2D-QMF) to separate the image into sub-images corresponding to different frequency bands. The decimated sub-images are then vector quantized using a trellis encoder. A trellis coder, rather than a codebook-based vector quantizer (VQ), is used since simulations have shown its advantage over VQ (both coders with random population) if the two coders are constrained to have the same amount of computations and same rate. Simulation results with images indicate that at 1 bit/pixel, the proposed system provides better subjective reconstructed image quality than the earlier reported SBC with scalar quantization.

1. INTRODUCTION

Subband Coding (SBC) is a medium complexity method for image coding at low bit rates. According to this method a filter-bank separates the original image into sub-images, each one representing a different frequency band. It is assumed that the image can be treated as a random process, with a slowly spatially varying spectral density. Then, if the frequency bands are sufficiently narrow, the sub-images, following decimation, can be considered as uncorrelated sources.

The separation to subbands can be performed using different two-dimensional filter-banks, e.g. The Laplacian pyramid [1], uniform filter-bank [2] and two-dimensional quadrature mirror filter-bank (2D-QMF) [3]. The filter-banks can be implemented in real time using fast convolvers which are already available commercially. The advantage of the 2D-QMF is that with no quantization, the analysis-synthesis system is approximately a unity system and complete cancellation of the aliasing, which results from the decimation process, is ensured.

In previously reported works [1, 2, 3] on subband coding of images, the subband signals were quantized by scalar quantizers. From rate-distortion theory we know that an improvement in performance should be obtained by using vector quantization, using sufficiently long vectors. There are three basic structures of vector coders [5]: Block (VQ), Tree, and Trellis coders. The block and tree coders are not instrumental for long vectors because their complexity grows exponentially with vector length (for a given coder rate). In this paper, we demonstrate that the trellis coder provides better performance than VQ when both coders are constrained to have the same amount of computations. Hence, we chose this coder for the SBC system we studied.

To use the trellis we have to populate it with codewords. To do this, two different methods may be used. In the first, a long training data sequence is used in an iterative process [8]. The second method is to populate the trellis from a random memoryless source having an appropriate probability distribution [4]. With this method one avoids the complicated and lengthy training process. However, the performance is dependent on how does the real image data approximate the assumed statistical properties.

The remainder of this paper describes in more detail the proposed coding algorithm and the results obtained. Section 2 introduces briefly the 2D-QMF filter-bank. Section 3 presents the trellis coder structure, the random population process, and simulation results (with synthetic data), from which we conclude that indeed the trellis is beneficial over VQ under the constraint of a given number of computations and the range of parameters examined. In Section 4 the coding system, including implementation considerations are presented. Simulation results with actual image data are given in Section 5. A summary and conclusions are contained in the last section.

2. 2D-QMF FILTER BANK

The basic structure of a 2D-QMF filter-bank is depicted in Fig. 1. It is a four-band array composed of analysis and synthesis filter-banks [3]. The outputs of the analysis stage are four sub-images, each of which is a quarter in size of the original image, representing different frequency regions, as illustrated in Figure 2.

To achieve complete cancellation of the aliasing due to decimation of the sub-images, it is common to select the filters which are determined from a single prototype filter, $H(w_1, w_2)$. The relations satisfied by the analysis and synthesis filters in order to obtain complete cancellation of aliasing, and the condition for a unity system are given in [3].

Using a separable prototype filter (i.e.: $H(w_1, w_2) = H_1(w_1)H_2(w_2)$) reduces dramatically the amount of computations needed, and enables one to use QMF filters designed for the one-dimensional case [7].

A filter-bank with more bands can be built by using the above basic 4-band 2D-QMF in a tree structure. The advantage of the tree-structured filter-bank is that it offers the flexibility of sub-dividing frequency regions of interest (where the input spectral density varies rapidly), while leaving other bands undivided.

3. TRELLIS VECTOR CODER

The trellis diagram can be described as a coding tree in which branches are joined. Consider, therefore, a coding tree of L levels (L is the tree depth). The first level consists of one node that splits into q branches. Each branch leads to a node which splits again into q branches, and so on, up to the last level. Each branch is populated with n codewords. In the coding process, a vector of $N=nL$ source elements is coded. The encoder looks for the path in the tree which will minimize the error between the source vector and the reproduction vector on the path. This path is transmitted to the decoder which has knowledge of the coding tree and reconstructs the reproduction vector from the transmitted path.

As stated above the trellis stems from the tree by joining branches. The joined branches are those for which the last k elements in the path description, leading to a given node, are identical. As a result, at each level (from the k-th on) there are q^{k-1} nodes and q^k branches. When all the levels are populated with the same codewords, the trellis is time invariant. On the other hand, if the codewords are not the same on each level, the trellis is time varying. In Figure 3 a time-invariant trellis in which $L=4$, $k=3$, $q=2$ is illustrated.

In this work, the trellis is populated using elements from a random process, with a probability distribution satisfying the rate-distortion function. The motivation for using random population is given in [4]. It is proven in [4] that, for a memoryless source, a time varying trellis populated randomly has performance (as $K \to \infty$), which is arbitrarily close to the rate-distortion function bound.

In the following, we will justify the use of a trellis code instead of VQ when random population is applied (assuming knowledge of statistical characteristics of the source). To determine which of the two coding structures is preferable, under the constraint that both have the same computational load (multiply-adds per source symbol), the following question is asked: Given the rate R, and the number of computations per source symbol $B = q^K$, how should one choose the trellis parameters K, n, and q, in order to get minimum distortion? The answer, based on simulation results, is demonstrated in table 1. The table shows the distortions obtained with different trellis parameters for $R=1$ bits/source-symbol, and for different values of B (64, 128 or 256). For the memoryless Gaussian source (with $\sigma^2 = 1$), the distortion criterion is mean square error (MSE), whereas for the memoryless Laplacian source ($\sigma^2 = 1$) it is the mean absolute error (MAE). In all the simulations, the length of the coded vector is $N = 256$.

The conclusion from these particular simulations is that with random population of a time varying trellis, under the constraint of a given number of computations, B, as above, the lowest distortion is obtained when the value of K is the largest possible (i.e. smallest q for a given B). The value of n is also the lowest possible which satisfies the rate constraint: $R = (\log_2 q)/n = 1$bits/symbol.

To compare between the performance of VQ and trellis coders, we used a VQ coder with vector size n' and codebook size N', set to satisfy the rate and computations constraints. As the trellis, the codebook is populated randomly. For the rate of 1 bit/symbol, and a given number of computations B, constrained as before, we have to set $N' = B$ and code vectors of length $n' = 6$, 7, or 8 samples, respectively. The results obtained now are identical to those obtained with the trellis for the particular choice of $K=1$. This case is depicted in Fig. 4. As seen in this figure, for $K=1$ the trellis is composed of L different codebooks in cascade. Thus, for each value of B, the distortion value shown in table 1, for $K=1$, is actually the average of the distortions obtained using L different random codebooks, each with same vector length.

Thus, we conclude that by using the proper trellis parameters, better results can be obtained with the trellis coder, than with VQ, for the same computational effort. Although this is achieved at the price of a larger memory size ($Lnq^k = NB$ for the trellis, compared to $n'N' = n'B$ for VQ).

4. SYSTEM DESCRIPTION AND IMPLEMENTATION CONSIDERATIONS

4.1 QMF-Array

For efficient implementation of the QMF array we have chosen the prototype filter $H(w_1, w_2)$ to be separable. Furthermore, identical one dimensional filters are used for each dimension: $H(w_1, w_2) = H(w_1) \cdot H(w_2)$

Among all 1D-QMF filters given in [7], the filters for each level were chosen such that they would give the best approximation to a unity system.

The 2D-QMF arrays used are full trees having either two or three levels.

To avoid the increase in the total number of output samples, due to the convolution operations, a periodic extension of the source image is introduced.

4.2 Coders description

A. Bit Allocation

The bit allocation between bands assures that for a given average rate, the MSE between the source image and the reconstructed one, D_T, is minimized. Assuming that there is a good band separation, we achieve this goal by minimizing the accumulated distortion of all the bands, since:

$$D_T = \sum_i D_i \tag{1}$$

where D_i is the distortion resulting from coding the i-th subband signal.

To find the optimal bit allocation, it is assumed that the coded image is a Gaussian source, and that the spectral density in each of the decimated bands is nearly constant.

With these assumptions, the optimal bit allocation is parametrically given by [9]:

$$r_i^* = \max\left(\frac{1}{2}\log\frac{\sigma_i^2}{\theta}, 0\right) \tag{2}$$

r_i^* – is the optimal rate given to band i (in bits/pixel)
σ_i^2 – the variance of the i-th band signal
θ – a parameter which can be calculated by solving the rate constraint, i.e. by forcing the total rate to be R,

$$R = \frac{1}{m}\sum_{i=1}^{m} r_i^* = \frac{1}{m}\sum_{i=1}^{m} \max\left(\frac{1}{2}\log\frac{\sigma_i^2}{\theta}, 0\right) \tag{3}$$

where m is number of bands.

Note that it is not possible to implement any desired rate by using the trellis, since q and n are integers. Choosing q to have the same value in all the bands, the possible rates are $r = (\log_2 q)/n$, $n = 1,2,...$.

Based on the optimal bit allocation, according to (2), n_i (the number of codewords on each branch of the trellis for band i) is determined to be the maximal integer such that:

$$r_i^* \le \frac{\log_2 q}{n_i} = r_i \tag{4}$$

where r_i is the actual rate given to band i. Therefore, the total rate is usually somewhat larger than the desired rate, R.

Typically, the lowest frequency region is characterized by high correlation between adjacent pixels, even after the decimation, a scalar DPCM coder is used in this band.

B. Trellis Population

Since we assumed that the original image data is Gaussian, it can be expected that the resulting decimated band signals are nearly memoryless Gaussian sources. Hence, the trellis coders (one for each band) can be populated randomly from a probability distribution which satisfies the rate-distortion function. On the other hand, looking at the histograms of different bands shows that the band distribution (apart from the lowest band) is very close to a Laplacian distribution. The problem is that a closed-form expression for the rate-distortion function for this distribution with an MSE distortion measure is not known. A closed-form expression is known only for the MAE distortion measure. But this measure is not suitable here since it does not satisfy (1).

In an effort to find a sub-optimal solution, two approaches were considered and examined. In both, the bit allocation was performed as described in the former section (under the Gaussian source assumption). The two differ by the distribution in which the trellis diagrams are populated. In the first, each trellis is randomly populated using the Laplacian distribution, with the variances being estimated from the data in each band (and sent as side information to the decoder).

In the second, each trellis was populated using the distribution satisfying the rate-distortion function for a Laplacian source (with the MAE distortion measure), given by

$$P(Y)=\frac{2\theta^2}{\sigma^2}\,\delta(y)+\frac{1}{\sqrt{2}\sigma}\,(1-\frac{2\theta^2}{\sigma^2})\,e^{\sqrt{2}|y|/\sigma} \qquad (5)$$

where θ is obtained from the bit allocation procedure described earlier, and σ is estimated for each band as before and is sent to the decoder. For both approaches, the search of a trellis is done under the MSE distortion measure. For rates of 1 bit/pixel and less, the second approach led to slightly better results. Therefore, in the proposed system the population of the trellis diagrams is done according to this second approach.

The memory size needed to store codewords of a time varying trellis is nLq^K. The trellis parameters used in the system ($K=3, q=32$) led to an unrealistic memory size requirement. This problem was overcome by using a table containing a much smaller number of random codewords (we stored 2500 words) and cyclically using them to populate each trellis. Using this method, a minor degradation in performance of only 0.1dB was observed.

C. Trellis Search

A full search of the trellis using the Viterbi algorithm [6] demands too much computational effort in large trellis structures. Thus, only a partial search is done using the M-algorithm [5], which is a sub-optimal version of the Viterbi algorithm.

Note, however, that already in the first level the trellis search begins at M nodes and, therefore, there is a need to send to the decoder side information about the node from which the selected path begins (($\log_2 M)/Ln$ bits/symbol). This information becomes negligible as longer vectors ($N = Ln$) are coded in the trellis.

4.3 In-Band Trellis Adaptation

The fact that the image is non-stationary, so that at different regions of the image its energy may concentrate in different frequency bands, forces the adjustment of the bit-allocation and the trellis population according to the spatial variations of the variance in each sub-image. Thus, instead of transmitting the variances and averages of the whole sub-image corresponding to a given frequency band, we need to transmit the variances and averages of sub-regions in each of the sub-images.

The regions in each sub-image are formed by dividing the sub image into equal size rectangles. For a 2-level filter-bank the sub-images were divided into 16 equal size regions. Whereas, for a 3-level filter-bank, the sub-images were divided only into 4 regions (since the decimated sub-images are smaller). Hence, for each of the above two filter-banks we get a total of 256 regions. To minimize the MSE, the bit allocation is now performed as in Section 4.2.A, but the variances of all 256 regions are taken into account. The trellis corresponding to each region is populated according to its local variance and the same parameter θ obtained from the bit allocation process. The variances and averages are coded by a uniform scalar quantizer.

5. SIMULATION RESULTS

In order to simulate the proposed system we used a Vax/750 computer with a Gould IP 8500 image display system. The simulations were performed on the image 'GIRL' of size 256×256 pixels, shown in picture 1.

In all trellis diagrams used, the coded vector size, $n_i L$, is 256. The branching factor, q, is the same in all bands. Its value was chosen to be 32 in order to enable rates up to 5 bits/pixel. The constraint length, K, is chosen to be 3 and the search parameter M, is 30. In all versions of the proposed algorithm the sub-image corresponding to the lowest frequency band is coded using a scalar DPCM coder.

When the system with adaptation was implemented at the rate of 1 bit/pixel, the following results were obtained: In the case where the 2D-QMF array contained 2 stages and the filter lengths were 32 at the first stage and 16 at the second, the SNR ($10 \log 255^2$/MSE) between the original image (Picture 1) and the reconstructed one (picture 2) was found to be 31.83 dB. Whereas, using a 3 stage 2D-QMF (with filter lengths of 32, 16, and 8), the SNR obtained is 31.48 dB (picture 3). Still the subjective quality of the reconstructed image in picture 3 is judged to be better.

SUMMARY AND CONCLUSIONS

The proposed system which consists of 2D-QMF and a trellis coder was found to achieve good results at the rates of 1 bit/pixel (for an image of 256×256). Although similar SNR's are obtained with a DPCM coder with adaptation [3], our experience has shown the trellis coded images to be subjectively better. Further improvement can be obtained by using more carefully selected parameters, and by populating the trellis according to better matched probability densities. Another possibility is to use a sufficiently long training sequence and the iterative process in [8].

We observed that some of the subband signals are not uncorrelated as assumed, and hence we are presently considering the application of a predictive trellis structure [10].

REFERENCES

[1] P.S. Burt and E.H. Adelson, "The Laplacian pyramid as a compact image code", IEEE Trans. on Communications, Vol. COM-31, No. 4, April 1983, pp. 532-540.

[2] B.L. Hinman, J.G. Bernstein and D.H. Staelis, "Short space Fourier transform image processing", ICASSP April 1984, pp. 4.8.1-4.8.4.

[3] J.W. Woods and S.P. O'Neil, "Subband coding of images", IEEE Trans. on ASSP, Vol. ASSP-34, No. 5, October 1986, pp. 1278-1288.

[4] A.J. Viterbi and J.K. Omura, "Trellis encoding of memoryless discrete-time sources with a fidelity criterion", IEEE Trans. on Inform. Theory, Vol. IT-20, No. 3, May 1974, pp. 325-332.

[5] N.S. Jayant and P. Noll, Digital coding of waveforms, Prentice Hall, 1984, pp. 428-461.

[6] G. D. Forney, "The Viterbi algorithm", Proceedings of IEEE, Vol. 61, No. 3, March 1973, pp. 268-278.

[7] E.C. Crochiere and L.R. Rabiner, Multirate digital signal processing, Prentice Hall 1984, pp. 401-404.

[8] L. C. Stewart , R.M. Gray and Y. Linde, "The design of trellis waveform coders", IEEE Trans. on Communications, Vol. COM-30, No. 4, April 1982, pp. 702-710.

[9] A. Segall, "Bit Allocation and Encoding for Vector Sources", IEEE Trans. on Inform. Theory, Vol. IT-22, No. 2, March 1976, pp. 162-169.

[10] E.D Ayanogly and R.M. Gray, "The design of predictive trellis waveform coders using the generalized Lloyd algorithm", IEEE Trans. on Communications, Vol. COM-34, No. 11, Nov. 1986, pp. 1073-1080.

TABLE 1: Distortion obtained for different trellis parameters

B	*K*	*q*	*n*	Gaussian	Laplacian
64	1	64	6	0.399	0.459
64	2	8	3	0.3559	0.436
64	3	4	2	0.353	0.432
64	6	2	1	0.347	0.429
128	1	128	7	0.377	0.447
128	7	2	1	0.3233	0.420
256	1	256	8	0.36	0.438
256	2	16	4	0.32	0.412
256	4	4	2	0.314	0.407
256	8	2	1	0.311	0.406

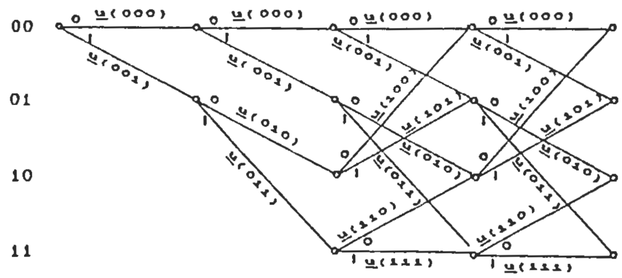

Figure 3. Trellis diagram.

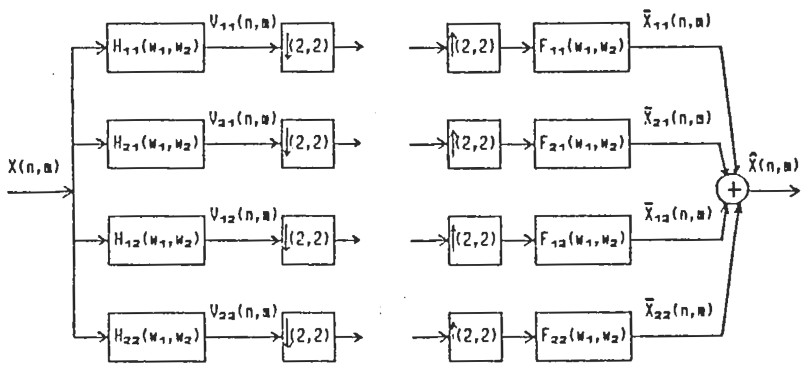

Figure 1. One stage of 2D-QMF.

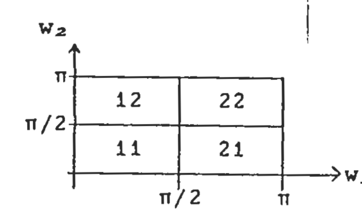

Figure 2. Subbands of 2D-QMF.

Picture 3.

Picture 2.

Picture 1.