NASA-CR-193729

IN-82-CR
180875
P. 10

# Low Bit Rate Coding of Earth Science Images[1]

Faouzi Kossentini, Wilson C. Chung, and Mark J. T. Smith
Georgia Institute of Technology
School of Electrical Engineering
Atlanta, Georgia 30332

## 1 Introduction

As the trend toward acquisition and storage of earth science imagery increases, the availability of good compression and coding techniques becomes more and more important. The particular compression needs among the earth science community are somewhat diverse. For some classes of images, there is a reluctance to accept any coding distortion. This can severely limit the amount of compression that can be achieved. In other cases, restrictions are milder and some amount of distortion can be tolerated as long as the visual results are perceptually indistinguishable from the original. There are also cases in which small but visible forms of distortion may be acceptable as long as the nature of the distortion does not interfere with the features of interest. In addition to quality, other system-related features may be important. For example, fast decompression capability for browsing archival databases, fast encoding capability for real-time transmission, and reduced sensitivity to channel coding errors for applications requiring transmission through noisy channels are sometimes high priority features. Given this variance in compression needs, a variety of compression approaches should be considered to address individual requirements.

In this paper, we discuss compression based on some new ideas in vector quantization and their incorporation in a subband coding framework. Several variations are considered, which collectively address many of the individual compression needs within the earth science community. The approach taken in this work is based on some recent advances in the area of variable rate *residual vector quantization* (RVQ). This new RVQ method is considered separately, (in the next section) and in conjunction with subband image decomposition (in the following section). Very good results were achieved in coding a variety of earth science images. The last section of the paper provides some comparisons that illustrate the improvement in performance attributable to this approach relative to the JPEG coding standard. Due to page limitations in these proceedings, only one earth science image is presented for illustration. However, a variety of coded earth science image examples will be presented at the conference.

# 2  Residual Vector Quantization (RVQ)

Residual Vector Quantization (RVQ), which is also called multistage vector quantization in the literature, consists of a cascade of VQ stages, each operating on the "residual" of the previous stage. Typically an RVQ will have many stages which can be indexed $i = 1, 2, \ldots, P$. Each stage codebook contains $N_i$ code vectors. Thus the total number of vectors in the RVQ is $\sum_{i=1}^{P} N_i$; however, the number of *direct sum* code vectors uniquely representable by the RVQ is $N = \prod_{i=1}^{P} N_i$. Hence the RVQ requires only a very modest amount of memory while being able to effectively represent a tremendous number of distinct code vectors. Each of these distinct vectors or direct-sum code vectors is obtained by adding the appropriate code vectors from each stage, hence the name direct-sum. The codeword associated with each direct-sum code vector is simply the concatenation of appropriate indices from each stage. Consequently, each direct-sum code vector is defined by one of $N$ indices, each of the same length.

Fixed rate RVQ is always upper bounded in performance by (conventional exhaustive search) vector quantization. The observed RVQ degradation in performance is believed to result from entanglements in the RVQ tree and from the direct-sum codebook structure. Thus it is important to design the RVQ carefully to obtain the best possible performance.

First, it has been shown in [2] that better performance can be achieved by designing the RVQ in a way that *jointly* optimizes each stage codebook. Necessary conditions for optimal fixed rate RVQ have been derived [2], and a jointly optimal RVQ design algorithm has been used [2, 11] to design fixed rate RVQ codebooks. While one of the conditions (the nearest-neighbor rule) suggests that all direct-sum vectors in the RVQ codebook are exhaustively searched (which is computationally demanding), earlier RVQ implementations were based on sequential search [2]. However, this fast searching technique leads to substantially reduced performance, as was observed in [2]. Multipath searching (introduced in [8] for tree searching) was shown [2, 9] to be very effective in trading a substantial reduction in computations for a small loss in performance. Interestingly the value of $M$ (number of searching paths) does not have to be very large in order to achieve performance close to that of exhaustive search. In particular, values of $M$ as low as 2 and 4 lead to significant improvement in performance. Therefore, an $M$-search RVQ design algorithm usually achieve a good compromise between performance and complexity.

## 2.1  Variable Rate RVQ

When the codewords are allowed to be variable, the average bit rate for the coded image can most often be reduced below that of a fixed rate system. Variable rate schemes usually exploit the entropy of the input signal thereby allowing a reduction in rate without a further reduction in quality.

The multistage structure of RVQ can be exploited, leading to a variable rate quantizer where the input vectors are encoded using different numbers of stages [10].

For each input vector, the encoding terminates once the distortion falls below a pre-determined threshold. Then, the number of stages used to encode an input vector is sent to the decoder as side information, which can be very costly. However, when relatively large vector sizes (such as $8 \times 8$ or $16 \times 16$) are used, side information requires only a small fraction of the total bit rate [10].

The variable rate RVQ as described above does not exploit the signal entropy explicitly. Another variable rate RVQ (described in [12, 14]) is based on a mathematical formulation that includes entropy explicitly as a parameter of the RVQ design problem. Very recently, necessary conditions for the optimality of variable rate residual vector quantizers were derived [12, 13], and a jointly optimal entropy-constrained RVQ (EC-RVQ) design algorithm (described next) was used to design locally optimal EC-RVQ codebooks.

## 2.2 The Entropy-Constrained RVQ Design Algorithm

The new entropy-constrained RVQ (EC-RVQ) design algorithm [12, 14] attempts to jointly optimize each stage codebook to minimize the reconstruction error over all training data subject to a constraint on the output entropy of the RVQ. It is an iterative descent algorithm based on a Lagrangian minimization (similar to the EC-VQ design algorithm [4]), where each iteration tries to simultaneously satisfy the necessary conditions given in [12, 13]. As in the design of optimal fixed rate RVQ codebooks, multipath searching is used in the encoder optimization step of the EC-RVQ design algorithm to closely satisfy the nearest-neighbor encoding rule. The $M$-search algorithm is found [14] to be very efficient in substantially reducing the encoding complexity of EC-RVQ for only a small loss in performance. Also in the design of EC-RVQ codebooks, the optimization of the decoder is the same as the one used in the design of fixed rate RVQ. The Gauss-Seidel algorithm (described in [13]) is used to find optimal stage code vectors (i.e. they simultaneously satisfy the centroid condition). Finally, the optimization of the lengths of the stage codewords can be done by incorporating an optimal coder design algorithm (such as the Huffman coding algorithm) into the EC-RVQ design algorithm, or by assuming an ideal entropy coder and using the *self-information* of a $P$-tuple index $j = (j_1, j_2, \ldots, j_P)$ as the optimal length of the *fictional* variable length codeword associated with that index $j$. It is shown in [12] that the self-information of a $P$-tuple index $j$ is also equal to the *sum* of $P$ stage conditional self-information components. Therefore, a sufficiently large training set can be used to find the conditional stage probabilities which will be used during the next iteration. The memory required to store all tables of stage probabilities can be very large, but a substantial reduction in memory requirements can be achieved if the lengths of the $p$th stage codewords are approximated or estimated by making a Markov-like assumption and using conditional probabilities which depend only on the last $m < p-1$ stages. As expected, the performance of EC-RVQ improves as $m$ approaches $P$. Empirical results in [11, 12] show that this improvement is very rapid and performance levels close to that of the optimal $P$th order EC-RVQ can be achieved with small values of $m$.

Entropy-constrained residual vector quantization or EC-RVQ has several important advantages over entropy-constrained VQ (EC-VQ) [3], [4] and the *alphabet and entropy constrained VQ* or AEC-VQ discussed in [16]. While EC-VQ has been shown to lead to a significant improvement in performance, it usually requires a large volume of memory for codebook storage and has a high level of computational complexity. A substantial reduction in memory could be made using AEC-VQ [16], which results in some small loss in performance but significantly less memory. However, the computational complexity still remains high. EC-RVQ is shown to be capable of achieving quality surpassing that of both AEC-VQ and EC-VQ, and its memory requirements are typically 30 times less than that of EC-VQ. Most important is that the computational requirements are only about 3% of that of EC-VQ and AEC-VQ.

The large computational complexity and memory requirements usually limit the codebook size used in EC-VQ and AEC-VQ, leading to variable rate VQs with rather limited output alphabet sizes. Therefore, the *peak* bit rate is usually not high (around 0.75 bits per pixel for 4 × 4 vectors), and EC-VQ and AEC-VQ fall short of achieving their performance potential. However, the EC-RVQ design algorithm can be used to design high rate codebooks, codebooks with high peak rates, and codebooks with relatively large vector sizes (such as 8×8). Since relatively high average encoding rates (typically 2.0 or 3.0 bits per pixel) are usually needed to encode low frequency image subbands, EC-RVQ can even be used to code these bands, resulting in a completely EC-RVQ based subband image coder. Indeed, subband/EC-RVQ systems work very well and are shown to achieve good quality at very low bit rates.

# 3 Coding Image Subbands Using EC-RVQ

The front-end of a subband image coding system (called the analysis system) decomposes the input into subband images by using maximally decimated filter banks. The dual process is called the synthesis system. It recombines the subbands using an interpolation filter bank. There are many factors that must be considered in designing and implementing the analysis/synthesis system, such as the frequency domain characteristics of the filter banks, the step response characteristics of the filters, the computational efficiency, the method for handling the image boundaries, and the subband decomposition structure. In this system we employ the computationally efficient IIR allpass polyphase filter banks based on two-band decompositions. A two-band systems is shown in Figure 2 and the corresponding polyphase structure is shown in Figure 3. The subband decomposition considered here is a rectangularly separable one in which the rows are first split into highpass and lowpass sequences and then the columns of the result are split in the same way. Four subband images result from this process, each with a distinct region of the frequency domain associated with it. In this implementation, the lowest frequency subband is further subdivided into four subbands producing a total of seven subband images. The recursive filter banks considered in this implementation are attractive from a computational efficiency viewpoint since they can be designed to have narrower transition regions than FIR filters of the same order of complexity.
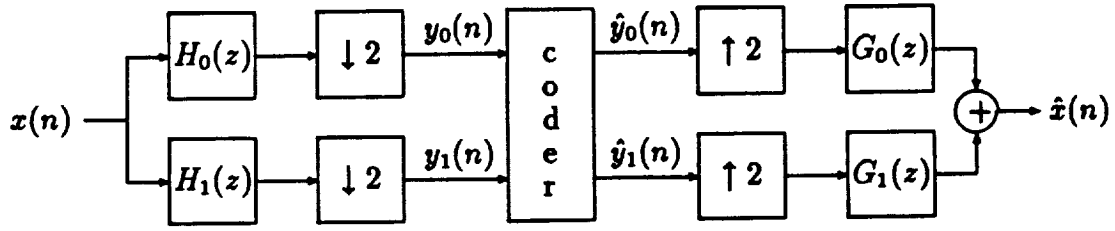
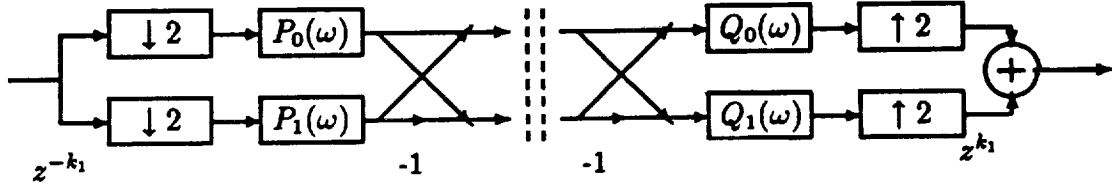Figure 1: Two-Band Analysis/Synthesis System



Figure 2: Two-Band Polyphase Implementation of the Analysis/Synthesis System

A few words are in order about the filters and their properties. The two-band system, shown in Figure 2, contains four filters: $H_0(z)$ and $G_0(z)$, the analysis and synthesis lowpass filters; and $H_1(z)$ and $G_1(z)$, the analysis and synthesis highpass filters. The analysis filters have the form

$$H_0(z) = P_0(z^2) + P_1(z^2)z^{-1} \tag{1}$$

$$H_1(z) = P_0(z^2) - P_1(z^2)z^{-1} \tag{2}$$

where

$$P_0(z) = \prod_{\ell=1}^{L} \frac{(1 - \alpha_\ell z^{-1})}{(z^{-1} - \alpha_\ell)} = \frac{\sum_{\ell=0}^{L} a_\ell z^{-\ell}}{\sum_{\ell=0}^{L} a_\ell z^{L-\ell}} \tag{3}$$

$$P_1(z) = \prod_{m=1}^{M} \frac{(1 - \beta_m z^{-1})}{(z^{-1} - \beta_m)} = \frac{\sum_{m=0}^{M} b_m z^{-m}}{\sum_{m=0}^{M} b_m z^{M-m}} \tag{4}$$

and $a_0 = b_0 = 1$. Ansari and Lui [1] have shown this class of filters to be extremely cost effective for sampling rate alteration applications. The high efficiency is due to the fact that each filter $P_0(z)$ and $P_1(z)$ is an allpass filter and symmetry can be exploited in the difference equation.

As discussed in [6] the set of synthesis filters for exact reconstruction are

$$G_0(z) = Q_0(z^2) + Q_1(z^2)z^{-1} \tag{5}$$

and

$$G_1(z) = Q_0(z^2) - Q_1(z^2)z^{-1} \qquad (6)$$

As before, the synthesis filters $G_0(z)$ and $G_1(z)$ may be implemented in the polyphase form shown in Figure 3 where $Q_0(z)$ and $Q_1(z)$ are called the synthesis polyphase filters. The polyphase synthesis section is exactly reconstructing when

$$Q_0(z) = \frac{1}{P_1(z)} \qquad (7)$$

and

$$Q_1(z) = \frac{1}{P_0(z)} \qquad (8)$$

In subband speech coding systems, (where input sequences may be viewed as being infinite in length), each channel is decimated to its respective Nyquist rate so that the number of input samples and analysis section output samples is essentially the same. Directly extending this idea to subband image coding systems (where the rows and columns are filtered as discussed), however, results in system which do not have this property. The sum of the pixels in the subband images is not equal to, but greater than the total number of pixels in the original. The problem in this 2-D case is that the image sizes are not infinite. Thus the convolution of the spatially limited image and the filter increases the size of the resulting image. The subband images, even though maximally decimated, have altogether a larger number of pixels than the original. To address this problem the method of circular convolution described in [6] was used to implement the systems. This method essentially treats the image as being periodically replicated. Filtering a periodic signal results in another periodic signal with the same period. It has been shown that by decimating and extracting one period of the result, data expansion can be avoided and exact reconstruction can be achieved [6].

In the next step, EC-RVQ is considered for coding the subband images. Entropy-based VQ has been considered previously for subband image coding and results were shown to be very good. In particular, Rao and Pearlman designed a set of codebooks using an AEC-VQ design algorithm [16]. In their work, only very small vectors (i.e. $1 \times 2$ and $2 \times 2$) were used for practical reasons. The use of larger vector sizes such as $4 \times 4$ requires a high computational demand and a large memory requirement (even though AEC-VQ requires the storage of only one codebook). However, the new EC-RVQ is very manageable in both regards and in fact allows for the use of larger vectors.

To achieve good coding performance, bit rates must be carefully allocated among the subbands. The average rate allocated to each of the $N$ subbands can be determined by minimizing the average distortion

$$D = \sum_{i=1}^{N} d_i(r_i) \qquad (9)$$

subject to the constraint that

$$\sum_{i=1}^{N} r_i \leq R \qquad (10)$$

**NODE**

$(b_{i,1}, D_{i,1})$

$(b_{3,2}, D_{3,2})$

$(b_{i,2}, D_{i,2})$

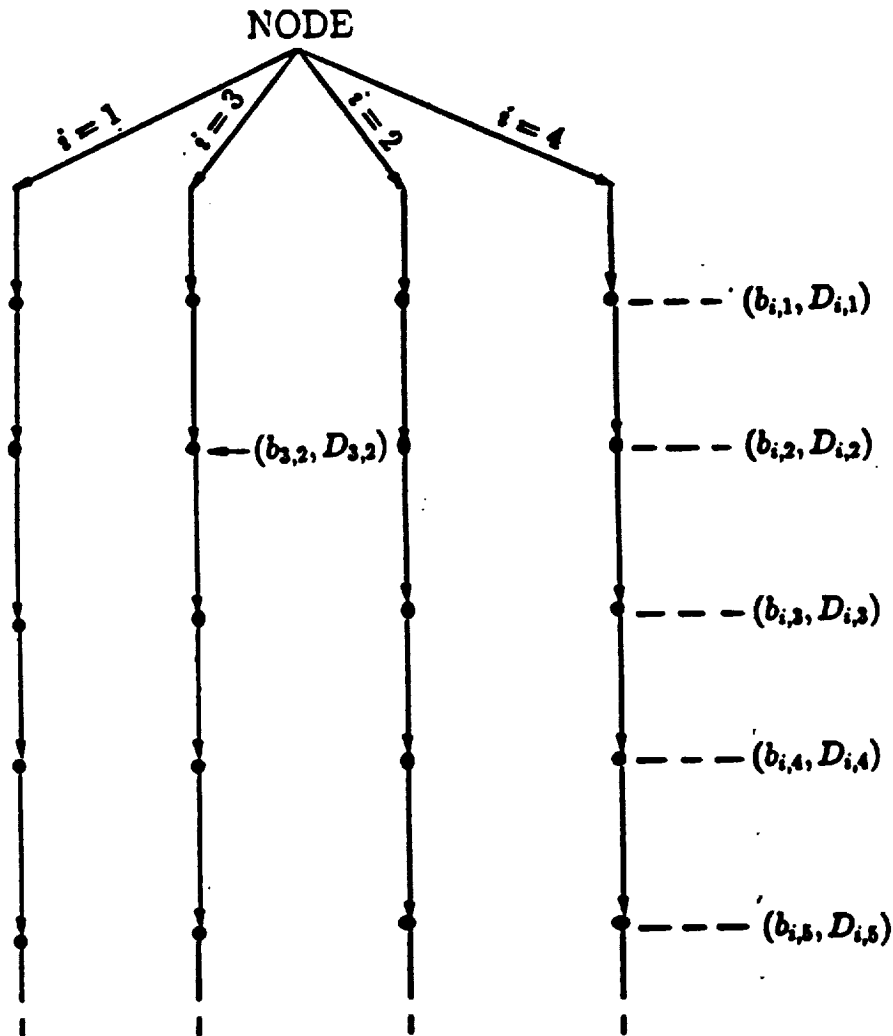$(b_{i,3}, D_{i,3})$

$(b_{i,4}, D_{i,4})$

$(b_{i,5}, D_{i,5})$

Figure 3: A four-band bit allocation tree

where $r_i$ is the average bit rate of the $i$th subband required to achieve a distortion $d_i(r_i)$, and $R$ is the maximum average bit rate for the entire image.

One approach to finding an optimal or near-optimal bit allocation is to determine the minimum of equation (9) by exhaustively searching all possibilities. In practice, however, the number of possible subband assignments may be astronomically large. Thus a more efficient solution is needed. Indeed, the algorithm derived in [15] can be used to efficiently allocate average rates among the subbands. In the procedure, a tree is constructed where the root node has $N$ children, one per subband, and the subtree rooted at each child $i$ is a unary tree of length $L_i$, where $L_i$ is the number of rate-distortion points available for the $i$th subband. Thus, each branch has $L_i$ nodes where each node $(i,j)$ $(1 \leq j \leq L_i)$ corresponds to a rate-distortion pair $(b_{i,j}, D_{i,j})$, where $b_{i,j}$ is the number of bits used to encode the $i$th subband by quantizer $j$ and

$D_{i,j}$ is the corresponding distortion such that (for $i = 1, 2, \ldots, N$)

$$b_{i,1} < b_{i,2} < \ldots < b_{i,L_i} \tag{11}$$

and

$$D_{i,1} \geq D_{i,2} \geq \ldots \geq D_{i,L_i}. \tag{12}$$

Therefore, the node $(i, 1)$ (the node closest to the root node) corresponds to the pair $(b_{i,1}, D_{i,1})$, and the node $(i, L_i)$ (the node farthest from the root node) corresponds to the pair $(b_{i,L_i}, D_{i,L_i})$. A four-band bit allocation tree is depicted in Figure 4.

Let $\mathcal{R}$ be a pruned subtree of the constructed tree $\mathcal{T}$, where the subband associated with the $i$th branch has now a rate $l_i$ and a distortion $\delta_i$. The average rate of the subband EC-RVQ associated with $\mathcal{R}$ is

$$l(\mathcal{R}) = \sum_{i=1}^{N} l_i \tag{13}$$

and the average distortion is

$$\delta(\mathcal{R}) = \sum_{i=1}^{N} \delta_i \tag{14}$$

where $b_{i,1} \leq l_i \leq b_{i,L_i}$ and $D_{i,1} \geq \delta_i \geq D_{i,L_i}$. The bit rate assignment problem can now be solved by finding the rates $l_1^*, l_2^*, \ldots, l_i^*$ that minimize $\delta(\mathcal{R})$ subject to $l(\mathcal{R}) \leq R$ ($R$ is the maximum average rate) over all pruned subtrees $\mathcal{R} \preceq \mathcal{T}$. The optimal pruning algorithm described in [15] gives such bit rates. Once the optimal average rate for each subband is determined, the encoder used for a particular subband is the one specified by the leaf node of the branch (associated with that subband) of the optimal tree $\mathcal{R}^*$.

# 4   Experimental Results and Comparisons

A variety of earth science images were considered in the training data set. In all cases, the test image was not included in the training set. Subband EC-RVQ codebooks were designed in the following way. First, a set of training images composed of a variety of earth science images was decomposed (as described in the previous section) into seven subband training sets. Vectors were extracted from the images within each subband to form seven sets of training data. These data were then used to train seven EC-RVQ coders. The peak bit rates ranged from 0.1 bits/pixel (bpp) to 4.00 bpp.

Figure 5a shows an original image of Moffett Field. Figure 5b shows the same image coded at 0.3994 bpp using the JPEG standard coding algorithm. For comparison, the same image is coded with the new subband/EC-RVQ coder at 0.4011 bpp. This image is shown in Figure 5c. The quality is better for the subband/EC-RVQ coder both objectively (28.05 dB for the JPEG image and 30.18 dB for the subband/EC-RVQ image) and subjectively (as seen in Figure 5). Our investigation thus far suggests that this coding approach has potential for improving the quality of compression of earth science images.

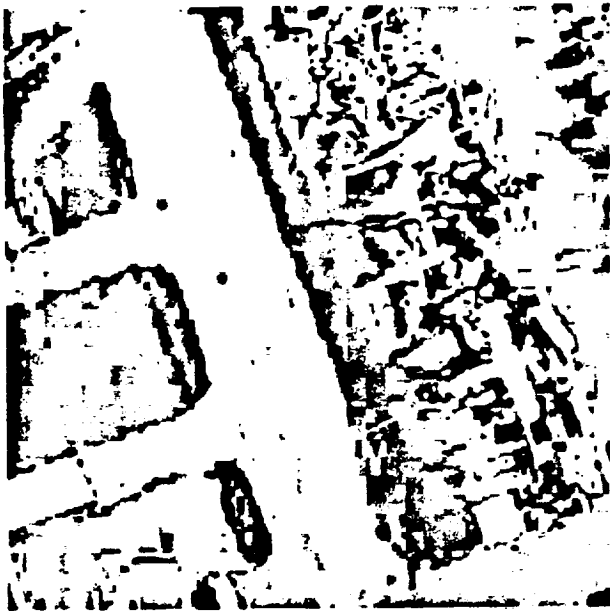Figure 5: a) Original Moffett image



Figure 5: b) Moffett image coded using JPEG Bit rate is 0.3994 bpp. PSNR is 28.05 dB.

Figure 5: c) Moffett image coded using subband/EC-RVQ Bit rate is 0.4011 bpp. PSNR is 30.18 dB

# References

[1] R. Ansari and B. Lui, "Efficient Sampling Rate Alteration Using Recursive (IIR) Digital Filters," *IEEE Trans. on ASSP*, Vol 33, pp. 90-97, 1985.

[2] C. F. Barnes, *Residual Quantizers*. PhD thesis, Brigham Young University, Provo, Utah, Dec. 1989.

[3] P. A. Chou, "Application of Entropy-Constrained Vector Quantization to Waveform Coding of Images," *VCIP IV*, SPIE-1199 pp. 970–978, 1989.

[4] P. A. Chou, T. Lookabaugh, and R. M. Gray, "Entropy-Constrained Vector Quantization", *IEEE Trans. on ASSP*, ASSP-37(1) pp. 31–42, January, 1989.

[5] J. C. Darragh, "Subband and Transform Coding of Images," Ph.D. Thesis, University of California, Los Angeles, 1989.

[6] S. L. Eddins and M. J. T. Smith, "Analysis/synthesis techniques for subband image coding," *IEEE Trans. on ASSP*, August 1991.

[7] J. H. Husoy, "Subband Coding of Still Images and Video," Ph.D. Thesis, University of Trondheim, Norway, 1991.

[8] F. Jelinek and J. Anderson, "Instrumentable tree encoding for information sources," IEEE Trans. on Info. Theory, IT-17 pp. 118-119, Jan 1971.

[9] F. Kossentini, M. Smith, and C. Barnes, "Large Block RVQ with Multipath Searching," ISCAS92, volume 1, pages 300–303.

[10] F. Kossentini, M. Smith, and C. Barnes, "Image Coding with Variable Rate RVQ," ICASSP92, March 23-26, 1992.

[11] F. Kossentini, M. Smith, and C. Barnes, " Finite-State Residual Vector Quantization," Submitted to *J. of Visual Com. and Image Rep.*, June 1992.

[12] F. Kossentini, M. Smith, and C. Barnes, " Locally Optimal Variable Rate RVQ," Submitted to *Trans. on SP*, January 1993.

[13] F. Kossentini, M. Smith, and C. Barnes, " Necessary Conditions for Optimal Variable Rate RVQ," Submitted to *Trans. on Info. Theory*, January 1993.

[14] F. Kossentini and C. Barnes, "Entropy-Constrained Residual Vector Quantization," Proceeding of ICASSP, Minneapolis, MN, April 27-30, 1993.

[15] E. A. Riskin, "Optimal bit allocation via the generated BFOS algorithm," *IEEE Trans. Info. Theory*, 37:400-402, March 1991.

[16] R. Padmanabha Rao and William A. Pearlman, "Multirate Vector Quantization of Image Pyramids," ICASSP91, volume 4, pages 2657–2660.

[17] P. H. Westerink, "Subband Coding of Images," Ph.D. Thesis, T. U. Delft, 1989.

[18] J. W. Woods and S. D. O'Neil, "Subband coding of images," *IEEE Trans. on ASSP*, ASSP-34, pp. 1278–1288, Oct. 1986.