# An Optimal Design Procedure for Intraband Vector Quantized Subband Coding

Jonathan N. Bradley, *Member, IEEE*, Thomas G. Stockham, Jr., *Fellow, IEEE*, and V. John Mathews, *Senior Member, IEEE*

*Abstract*—**Subband coding with vector quantization is addressed in this paper. Forming the data vectors from both between and within the subbands is considered. The former of these two schemes is referred to as interband coding and the latter as intraband coding. Interband coder design is relatively straightforward since the design of the single codebook involved follows readily from a representative set of interband data vectors. Intraband coder design is more complicated since it entails the selection of a vector dimension and a bit-rate for each subband. The main contribution of this work is an optimal methodology for intraband subband vector quantizer design. The problem formulation includes constraints on the bit-rate and the encoding complexity and is solved with nonlinear programming methods.**

**Subband vector quantization image coding in conjunction with a human visual system model is thoroughly investigated. Results of a large number of experiments indicate that the optimal intraband coder yields superior results from quantitative as well as subjective points of view than the interband coder for comparable bit-rates. This improvement becomes more pronounced as the computational complexity of the intraband encoder is allowed to increase.**

## I. Introduction

The closely related ideas of subband coding (SBC) and transform coding (TC) have existed for more than a decade [1], [2], [3]. These methods involve a decomposition of a data signal into a number of subsignals (referred to here as *subbands*), each of which represents information in a particular frequency band. The subbands are quantized individually and subsequently recombined. By coding the subbands separately, higher signal-to-noise ratios are obtainable than by single channel methods in general since the coding algorithm can be better matched to the signal statistics. In addition, since many perceptual criteria (both visual and auditory) are formulated in the frequency domain, these techniques allow for weighting the various frequency components so that the quantization noise can be matched to such criteria. Traditionally, pulse code modulation (PCM) or differential pulse code modulation (DPCM) is used for coding the individual subbands. Optimizing such a system involves an assignment of bits among the subbands so as to minimize the overall distortion while maintaining some pre-assigned overall bit-rate [4].

Paper approved by Arun N. Netravali, the Editor for Image Processing of the IEEE Communications Society. Manuscript received: January 29, 1991; revised: May 11, 1992. This work was supported in part by the National Science Foundation under grant no. MIPS-9016331.

J. N. Bradley is with Los Alamos National Laboratory, MS B-265, Los Alamos, NM 87545 USA.

T. G. Stockham, Jr. and V. J. Mathews are with the University of Utah, Department of Electrical Engineering, Salt Lake City, UT 84112 USA.

IEEE Log Number 9410085.

0090-6778/95$4.00 © 1995 IEEE

Vector quantization (VQ) is a coding technique that has attracted much attention lately [5], [6], [7]. Rather than quantizing individual data samples (as in PCM), VQ quantizes blocks of data samples. VQ achieves coding gain through the tighter sphere packing obtained in higher dimensions and, since it possesses memory, by exploiting statistical dependencies. This work examines the incorporation of VQ in SBC. The ideas introduced here are equally applicable to combining VQ with TC (and is explored in [8]).

This paper is concerned with vector quantization of subband decompositions of signals. When designing a subband vector quantizer (SBVQ), one has two fundamental options as to how to form the data vectors. One alternative is to form the data vectors as slices from across the subbands, i.e., with each data vector containing at most one data sample from any given subband. We will refer to such schemes as *interband coding*. The design of such a quantizer is relatively straightforward since no bit allocation procedure is involved. The emphasis placed on coding the different subbands is embodied in the codebook which is designed from a representative set of interband data vectors. The other alternative is *intraband coding* in which the data vectors are formed from within each subband. Here the subbands are coded separately, each with a unique codebook. The design of an intraband SBVQ system involves the selection of a vector dimension and codebook size (or equivalently, a vector dimension and a bit-rate) for each subband.

It should be noted that when the entire spectral coefficient vector is encoded as a single vector, as is considered in this work, the system is not capable of exploiting all of the degrees of freedom offered by subband coding. In fact, in the case when the subband decomposition is orthogonal and the filter length is less than or equal to the downsampling factor, such a scheme is equivalent to performing straight VQ on the original data and no gain whatsoever is obtained from the subband decomposition. Interband SBVQ systems where the spectral coefficient vector is divided into smaller vectors and coded with separate vector quantizers (as is explored in [9]) is an alternative that more fully takes advantage of subband coding. However, although coding the spectral coefficient vector as a single vector is more computationally intensive, it gives a lower bound on the distortion achievable with other interband SBVQ methods.

Westerink, et al. [10] applied the interband concept to image coding through the use of a two-dimensional subband decomposition. A theoretical result relating the gain

of SBVQ to SBC/PCM was derived and experimental results relating the SNR of the SBVQ scheme to other well-known image coding methods were presented. Results of a very similar image coding system are presented in [11]. Applications of SBVQ in speech coding can be found in [9], [12], [13], [14], [15], [16]. In all of the works referenced that involve intraband SBVQ, the vector dimensions and codebook sizes of the vector quantizers were selected *a priori* and consequently were not optimal in any sense. The major contribution of this work is the development of a design methodology for selecting these parameters in an optimal fashion.

In [17], a method was discussed in which the subband quantizers could be optimally selected from a set of admissible quantizers such that a constraint on the overall bit rate is satisfied. The selection is based on computation involving the convex hull of the rate-distortion coordinates for the admissible quantizers. Reference [18] investigated two image subband coding systems in which the bit-allocation was accomplished by selecting the subband quantizers from an admissible set of entropy-coded uniform threshold quantizers. The effects of channel errors on the system performance were also studied in this work.

The novel design algorithm introduced in this work contrasts with previous methods in two significant ways: (1) each subband quantizer has a rate-distortion characteristic described by a convex parametric function and need not have a bit-rate selected from a prespecified finite set and (2) a constraint on the overall encoding complexity of the system is included. The second constraint is quite advantageous: selecting the vector quantizer parameters with regard only to bit rate can easily result in a system of unwieldy complexity.

The next section discusses the basics involved in implementing the subband analysis and synthesis systems. VQ is also reviewed—of particular interest is the discussion of the rate-distortion properties of this type of quantizer. Section III presents the intraband SBVQ design algorithm. A model of human vision and its application in image coding are discussed in Section IV. Section V contains experimental results from an application of the introduced ideas to monochromatic image coding. Experiments were performed with both optimal intraband SBVQ and an interband SBVQ system similar to that presented in [10], thus providing an evaluation of the relative merits of the interband and intraband concepts.

## II. Preliminaries

### *The Subband Coder*

A block diagram of an $M$-band intraband SBVQ system is shown in Figure 1 [4], [19]. It is assumed in this discussion that the analysis filters have equal bandwidth with square frequency support in the $\omega_1 - \omega_2$ plane and, consequently, that the downsamplers effect a decimation by a factor of $\sqrt{M}$ in each dimension. Input vectors are formed by appropriately sampling from the subband signals as discussed earlier and vector quantized. The indices of the selected codevectors are stored or transmitted. At

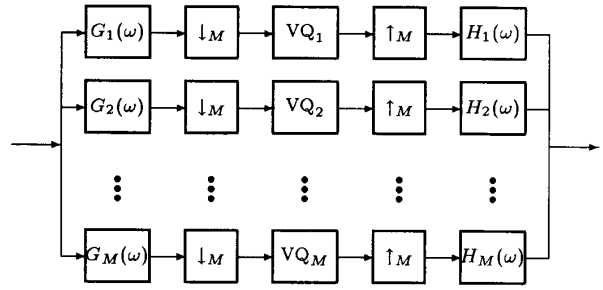


Fig. 1. Intraband SBVQ system.

the receiver, the corresponding codevectors are read from a look-up table to reconstruct the quantized subband signals. Each of these signals is then upsampled and bandpass filtered. The outputs of these filters (referred to as synthesis filters) are summed to yield the synthesized and quantized signal. The analysis and synthesis filters are designed such that no distortion is introduced by the combination of the analysis and synthesis processes in the absence of quantization. The design of such filters is discussed in [20].

### *Vector Quantization*

A vector quantizer [5], [6] consists of an encoder-decoder pair that is defined by an $N$ element codebook of $k$-dimensional vectors. The bit-rate of the quantizer (the number of bits required for transmission of each input sample) is

$$r = \frac{\log_2 N}{k}.$$

The encoder subdivides the discrete-time input signal into $k$-dimensional data vectors $\mathbf{x}_n$ and then (ideally) for each $\mathbf{x}_n$ determines that codevector $\mathbf{y}_i$ that is most representative. This is equivalent to determining $\mathbf{y}_i$ such that

$$d(\mathbf{x}_n, \mathbf{y}_i) \leq d(\mathbf{x}_n, \mathbf{y}_j) \quad \forall j \neq i$$

where $d(\mathbf{x}, \mathbf{y})$ denotes some measure of distortion between two vectors. When the mean-square error (MSE) between the input and quantized signals is of interest, the appropriate distance measure is the squared Euclidean distance given by

$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{k} (x_i - y_i)^2. \tag{1}$$

A straightforward method for implementing the encoder is to compute the distance between the input vector and each element of the codebook and then select that codevector corresponding to the minimum of these distances. For the distance measure (1) this entails $N$ distance computations which corresponds to approximately $Nk$ multiplications and $2Nk$ additions for each input vector. Denoting the complexity of performing two adds and one multiply as $\alpha$ and assuming that the cost of performing the comparisons is negligible, the complexity $Q$ of encoding one data

sample in an unstructured vector quantizer is then [1]

$$Q \approx 2^{rk}\alpha. \tag{2}$$

The decoder has knowledge of the same codebook used by the encoder and constructs the quantized signal from the received codevector indices. The implementation of the decoder consists simply of a table look-up, the complexity of which is trivial compared to that of the encoder. It is the encoding complexity that constrains the selection of $k$ and $N$. The codebook is designed from a set of training data so as to minimize the expected value of the quantization distortion. Several iterative techniques are available for performing this task: see, e.g., [6], [21].

An accurate description of how the distortion due to vector quantization relates to bit-rate is necessary in this work since the task of designing a family of vector quantizers for the subbands requires a knowledge of how allocating resources among the various subbands affects the overall distortion in the quantized image. In [22] the Euclidean rate-distortion characteristic of an optimal vector quantizer was derived as

$$D(r) \propto 2^{-2r} \tag{3}$$

where the constant of proportionality is dependent upon the vector dimension and the multidimensional probability density function of the data vectors. However, the derivation was based on the assumption of large $N$, i.e, "high quality" quantization. Since the compression ratios that are explored here for the intraband SBVQ system are on the order of 0.25–0.5 bits/pixel, the individual subband quantizers do not often operate in the high quality region and it is necessary to model rate-distortion performance by expressions different from those given by asymptotic theories. Figure 2 depicts experimental rate-distortion measurements of vector quantized image subband data displayed on a dB scale along with the corresponding linear regression curves. These data were obtained by decomposing a set of 512 x 512 eight-bit gray scale images with the sixteen-band decomposition discussed in Section V and then measuring the distortion that was obtained by vector quantizing one of the 128 x 128 subbands with 2- and 4-dimensional vector quantizers of various rates. This diagram suggests that the distortion-rate characteristic for vector quantization of image subband data is approximately exponential. However, the two linear distortion curves have slopes of -8.06 and -10.00 dB/bpp, values that depart significantly from the slope of -6.02 dB/bpp as predicted by (3). Based on such empirical observations (which were observed for the other subbands as well) we have modeled the mean-square quantization error $D(k, r)$ for VQ using the formula

$$D(k, R) = \beta(k)e^{-\gamma(k)R} \tag{4}$$

where the constants $\beta(k)$ and $\gamma(k)$ depend upon the vector dimension and must be determined experimentally. In this paper these parameters are estimated from the slope and the intercept of linear regression curves such as those in Figure 2.

[1] If the squared-norms of the codevectors were precomputed, $\alpha$ would then be the complexity of performing one multiplication and one addition.

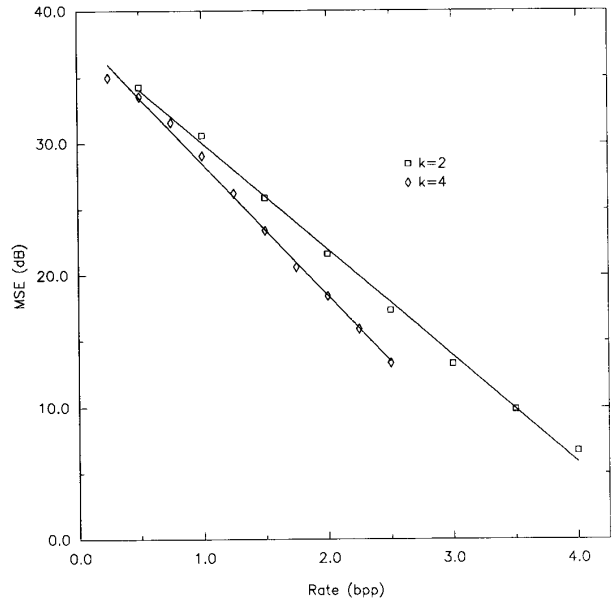


Fig. 2. VQ distortion-rate measurements for subband 10-10. (See Figure 5 for a depiction of the subband support).

## III. Intraband SBVQ System Design

The problem of designing an intraband subband vector quantizer involves the selection of the vector dimension and codebook size for each subband. Once these parameters are specified, each codebook can be obtained by applying the generalized Lloyd algorithm [6] (or another suitable clustering algorithm) to an appropriate set of training data. The method for selecting these parameters involves an extension of the reasoning used in the approach to vector source coding [23] that has been widely applied to subband and transform coding [4]. In [23], the bit-rate of the coder for each vector component was obtained via an optimization problem which specified that the bit-rates be selected so as to minimize the overall distortion of the coding scheme while maintaining some prespecified value for the overall bit-rate. This technique is commonly applied when the individual coders are PCM or DPCM. However, for the application under consideration here, a different approach must be taken. For a given bit-rate, lower distortion can be obtained with VQ by increasing the vector dimension. Thus, if only a rate constraint is imposed, the optimization algorithm will select an arbitrarily large vector dimension for each vector quantizer. For this reason, a constraint on the overall encoding complexity was included in the problem formulation.

The design problem of intraband SBVQ for the full-search case is discussed here in detail. The system consists of $M$ equal-bandwidth subbands and is to operate at a maximum rate of $R$ and complexity less than or equal

to $Q$. Let the vector dimension and bit-rate of the vector quantizer for the $i$-th subband be, respectively, $k_i$ and $r_i$. Then, the number of codevectors for the $i$-th subband subband is given by

$$N_i = 2^{r_i k_i}.$$

Before the optimization can proceed, the quantities that parameterize the VQ rate-distortion behavior in each subband must be estimated. Since these estimates are made empirically, the vector dimension in the $i$-th subband must be restricted to belong to some prespecified set $K_i$. The intraband SBVQ design problem is formulated as:

Minimize

$$D = \sum_{i=1}^{M} \beta_i(k_i) e^{-\gamma_i(k_i) r_i} \tag{5}$$

over the $k_i$ and $r_i$ subject to

$$\frac{1}{M} \sum_{i=1}^{M} r_i \leq R \tag{6}$$

$$\frac{1}{M} \sum_{i=1}^{M} 2^{r_i k_i} \alpha \leq Q \tag{7}$$

$$r_i \geq 0 \tag{8}$$

$$k_i \in K_i. \tag{9}$$

The objective function given by equation (5) represents an expression for the overall distortion introduced to the signal by the subband vector quantizer. The basic assumptions employed here are that the Euclidean distance measure is used and that the various subbands are mutually orthogonal to each other. Consequently, the total distortion is the average of the distortions incurred in each subband. The rate constraint (6) is derived from the fact that the overall bit-rate of the system is the average of the bit-rates of the individual vector quantizers. Similarly, the complexity constraint (7) states that the complexity of encoding each pixel is the average of the encoding complexity in each subband. The $r_i$ are necessarily discrete since a vector quantizer must have an integral number of codevectors. However, for simplicity, it is assumed that the $r_i$ are continuous and after a solution is found, the $N_i$ are calculated by

$$N_i = [2^{r_i k_i}] \tag{10}$$

where $[\cdot]$ denotes the function that maps a real number to the closest integer. It is impractical, however, to treat the $k_i$ as continuous.

A significant aspect of this optimization problem is that it contains unknowns of both continuous and discrete nature: the $r_i$ are continuous whereas the $k_i$ must be selected from a finite set of integers. If the $k_i$ are fixed, the optimization of the $r_i$ is a well-understood exercise in nonlinear optimization. However, the optimization of the $k_i$ involves integer programming. Consequently, the overall problem is solved by an ad hoc procedure in which the problem is decomposed into two subproblems, each of which is concerned with the optimization of one of the two classes of unknowns. This two-stage solution was chosen over a joint optimization of the variables in order that information about the gradient of the objective function with respect to the $r_i$ could be utilized.

In the first subproblem, which is concerned with the selection of the $r_i$, the $k_i$ are held constant and (5) is minimized over the $r_i$ subject to the constraints (6), (7), and (8). Let $\mathbf{k}$ denote an $M$-dimensional vector whose $i$-th entry is $k_i$. Also, let us denote the minimum distortion for a given $\mathbf{k}$ as $D^*(\mathbf{k}, R, Q)$. This problem involves the minimization of a convex objective function over a convex constraint set and hence possesses a unique local minimum. Performing this optimization requires a nonlinear programming technique due to the presence of the complexity constraint (7). In this work we employed a sequential quadratic programming technique to solve this subproblem. Details of the method can be found in [24].

The second subproblem is concerned with the selection of $\mathbf{k}$ which, due to (9), is restricted *a priori* to belong to a finite set. In theory the optimal $\mathbf{k}$ could be determined by evaluating $D^*(\mathbf{k}, R, Q)$ for each possible $\mathbf{k}$, however, this is typically an unacceptably large number of optimizations to perform. The selection of the $k_i$ is therefore based on a heuristic procedure that makes exploratory moves in $\mathbf{k}$ and checks for a decrease in $D^*(\mathbf{k}, R, Q)$. The search is described as follows. Assume that the elements of each $K_i$ are indexed in ascending order. Let $k_i$ (the $i$th element of $\mathbf{k}$) be indexed as the $j_i$th element of $K_i$, i.e. $k_i = K_i(j_i)$.

1. Select initial $j_i$, set $D = D^*(\mathbf{k}, R, Q)$.
2. Set $i = 1$.
3. (a) Let $k_i = K_i(j_i + 1)$, if $D^*(\mathbf{k}, R, Q) > D$ go to 4.
   (b) $D = D^*(\mathbf{k}, R, Q), j_i = j_i + 1$, go to 6.
4. (a) Let $k_i = K_i(j_i - 1)$, if $D^*(\mathbf{k}, R, Q) > D$ go to 5.
   (b) $D = D^*(\mathbf{k}, R, Q), j_i = j_i - 1$, go to 6.
5. $k_i = K_i(j_i)$.
6. $i = i + 1$, if $i \leq M$ go to 3.
7. If any $k_i$ changed in the above loop go to 2.
8. Exit.

The algorithm exits when, for any $i$, replacing $k_i$ by a neighboring element of $K_i$ results in an increase in $D^*(\mathbf{k}, R, Q)$. Such a selection of $\mathbf{k}$ can only be guaranteed to be locally optimal.

## IV. The Multiplicative Visual Model and its Application in Image Coding

In general, the problem of quantizer design is formulated so as to minimize some quantitative measure of distortion between the input image and the quantized result. The selection of the distortion measure is often motivated by the mathematical tractability of the resulting optimization problem and not necessarily by the subjective image quality resulting from its use. This section discusses a model of human vision and its use in improving the perceptual quality of an image coding system.

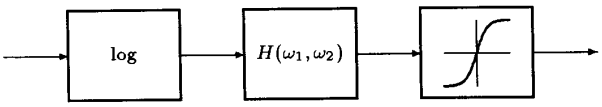


Fig. 3. The multiplicative visual model.

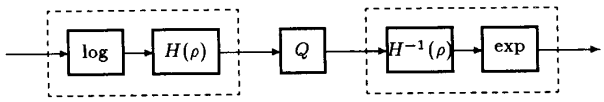


Fig. 4. Use of the multiplicative visual model in improving subjective quality of quantized images.

The multiplicative visual model (MVM), shown in Figure 3, was utilized in this work. It was first proposed by Stockham [25] and is intended to describe the early retinal stages of the visual process. The proposition of the model was partially based on a hypothesis of multiplicative image formation: an image was assumed to be composed of a product of a reflectance component and an illumination component. Reasoning that a system that processes images should obey superposition on the same arithmetic operation with which images are formed, Stockham argued that the visual system can be modeled as a multiplicative homomorphic system. Hence, as in the canonical form for the multiplicative homomorphic system [26], the visual model contains a logarithm in cascade with a linear system. The saturation element has been included to account for limitations of the photoreceptor outputs. The general nature of the linear component (denoted by $H$) can be inferred by considering how the visual system interprets a scene. Due to wide fluctuations in illumination, intensity images with dynamic ranges of 30 dB or more are often encountered. Despite the fact that different objects in an image may have greatly different illuminations, the human visual system accurately perceives their relative shades. This effect is explained by reasoning that the visual system acts to suppress the illumination. Since, in general, the illumination is a lowpass signal and the reflectance is highpass, suppression of the illumination could be accomplished if $H$ were to have a highpass characteristic. Such a choice for $H$ in a multiplicative homomorphic system results in an automatic gain control system [27]. The MVM explains several visual illusions [28] and also unifies three well-known (apparently conflicting) visual models developed on the basis of data obtained from psychophysical and electrophysiological experiments [29].

In determining a specific form for $H$ it was first reasoned that rotating an image should effect a similar rotation of the perceived image. This specifies $H$ to be "circularly symmetric," that is, $H(\omega_1, \omega_2) = H(\rho)$ where $\rho = \sqrt{(\omega_1^2 + \omega_2^2)}$. A more specific form for $H$ was deduced by surmising that $H$ must be "size invariant," i.e, that $H$ should process objects in an image similarly regardless of their distance from the viewer. The overall form of the filter is

$$H(\rho) = a\rho^{\mu} + b$$

where the specific values of $b$ and $\mu$ are selected via a trial-and-error process designed to adjust $H$ in order that $H^{-1}$ successfully reverse several visual illusions [28], [30]. The parameters used in this work are $b = 0.33$ and $\mu = 0.255$. The value of $a$ is such that $H(\rho)$ is equal to 2.0 at the Nyquist frequency.

The system in Figure 4 illustrates how the MVM is used to improve the subjective quality of an image coder. It was assumed that the dynamic range of the input image is always within the limits of the saturation element shown in Figure 3 and hence this system was ignored. The idea is to preprocess an image by the MVM in order to transform it to a "visual space." This "previsualized" image is then quantized by the quantizer $Q$ that is designed to minimize MSE *within* this space. The quantized visual domain image is then processed by the inverse of the MVM before it is viewed. It is hypothesized that minimizing MSE withing the visual space yields a quantizer better matched to human vision [28]. This method allows the use of the mathematically tractable MSE relationship in the quantizer optimization.

## V. Experimental Results

This section presents the results of a variety of experiments in which SBVQ systems were used to compress monochromatic still images of size 512×512 pixels at 8 bits/pixel resolution. Analyses were performed both with and without the MVM included. A two-dimensional sixteen-band quadrature mirror filter (QMF) bank [20] was used to perform the subband decomposition; a labeling convention for the subbands (equivalent to the one discussed in [20]) and their frequency support are depicted in Figure 5. Details about the QMFs used here are given in

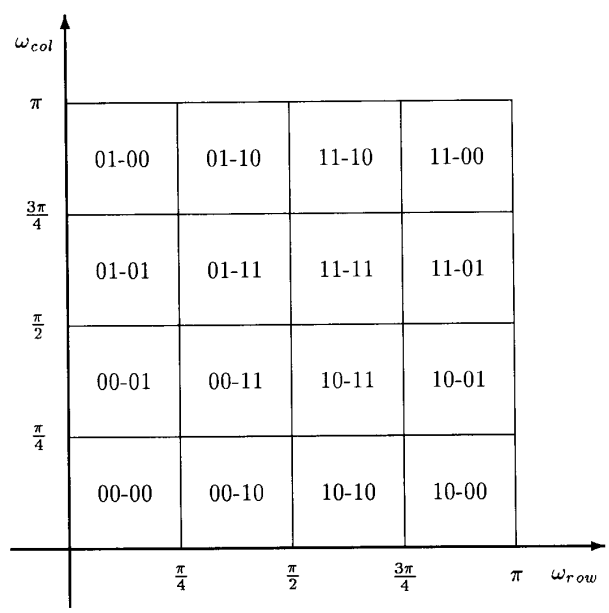


Fig. 5. Frequency support of subband decomposition.

Fig. 6. Original Lenna image.

Fig. 7. Lenna image coded by previsualized interband SBVQ with $R = \frac{9}{16}$.

[8]. Two fidelity measures were used to evaluate the effectiveness of the coding explored: for previsualized images signal-to-noise ratio (SNR) was used

$$\mathrm{SNR} = 10 \log_{10} \frac{\sigma_s^2}{\sigma_e^2}$$

and for for comparisons with raw image data, peak signal-to-noise ratio (PSNR) was used

$$\mathrm{PSNR} = 10 \log_{10} \frac{255^2}{\sigma_e^2}$$

where $\sigma_s^2$ is the signal variance and $\sigma_e^2$ is the variance of the quantization noise. No entropy coding was performed on the codevector indices.

The use of the visual model requires the implementation of a logarithm and an exponential (see Figure 4). In this work the raw image data were (as is most often the case) available in the "density" domain, i.e., the logarithm was inherent in the image recording process. Moreover, no computational resources were required for the exponentiation which was effected by gamma correction in the image display.

*Interband SBVQ*

Six interband SBVQ coders of various bit-rates were designed for previsualized image coding. The bit-rate varied from $\frac{1}{4}bpp$ to $\frac{9}{16}bpp$ in $\frac{1}{16}bpp$ increments, that is, 16-dimensional codebooks containing 16, 32, 64, 128, 256, and 512 codevectors were designed. Sixteen 512 x 512 training images were used in each codebook design. The results of coding the image in Figure 6 for $R = \frac{9}{16}bpp$ and $R = \frac{1}{4}bpp$ are shown in Figures 7 and 8, respectively.

*Intraband SBVQ*

The same sixteen image training set used in obtaining the interband coders were used to design a collection of codebooks of various rates and vector dimensions for each of the sixteen subbands. Four additional images from outside of this set were quantized to obtain the distortion-rate values used in estimating the $\beta_i(k_i)$ and $\gamma_i(k_i)$. The sets $K_i$ were selected as $1, 2, 4, 8, 16, 32, 64$ for each subband. It was assumed that image data are isotropic and therefore that the VQ rate-distortion parameters from subbands with frequency support symmetric about $\omega_1 = \omega_2$ are equivalent. Hence, it was only necessary to estimate the distortion-rate characteristics of ten subbands rather than sixteen.

In intraband SBVQ design, one has two degrees of freedom in specifying the system optimization problem (rate and encoding complexity) whereas in the interband case there is only one (rate). In a first attempt to contrast intraband SBVQ with interband SBVQ, a set of six intraband coders were designed having the same bit-rates and encoding complexities as in the interband experiment. The encoding complexity is given by $Q = 2^{16R}\alpha$ which is obtained from (2). The vector dimension and codebook size assignments calculated for the six cases are shown in Table 1. The results of coding the image in Figure 6 for $R = \frac{9}{16}bpp$ $Q = 512\alpha$ and for $R = \frac{1}{4}bpp$ $Q = 16\alpha$ are shown in Figures 9 and 10, respectively.

In order to explore the coding gain available from using intraband coders with larger complexities a second experiment was performed. A set of coders having the same bit-rates as those in the previous experiment was designed but with each coder having an encoding complexity of $512\alpha$.

Fig. 8. Lenna image coded by previsualized interband SBVQ with $R = \frac{1}{4}$.

Fig. 9. Lenna image coded by previsualized intraband SBVQ with $R = \frac{9}{16}$ and $Q = 2^{16R}\alpha$.

The coders corresponding to $R = \frac{9}{16} bpp$ were identical for the two experiments. The vector dimension and codebook size assignments calculated for the six cases are shown in Table 2. The quantized image for $R = \frac{1}{4} bpp$ $Q = 512\alpha$ is shown in Figure 11.

Figure 13 illustrates the rate-distortion performance of the coders designed in the above experiments. Each distortion value is obtained by averaging SNR measurements made within the visual domain on four images from outside of the training set. Although Figure 13 depicts the appropriate quantity for evaluating coder performance (since the quantizers were designed to minimize MSE within the MVM) Figure 14 is included which shows average PSNR measurements made on the raw image data. This is included to provide a comparison with results tabulated in the literature.

One may well question the validity of modeling the $r_i$ as continuous variables in the intraband problem formulation. For this reason, the actual complexity of each intraband coder is tabulated in Tables 1 and 2 for comparison with the design complexity. It is seen that rounding the $r_i$ as in (10) does not lead to a significant violation of the constraint. In some cases, the complexity of the resulting coder is significantly less than the design complexity. This is not due to rounding the $r_i$ but is rather a result of the limited number of allowable vector dimensions for each subband.

In Figure 13 it is seen that (at the expense of higher memory requirements) intraband SBVQ yields higher quality coded images than interband SBVQ for a given bit-rate and encoding complexity. The improvement becomes more pronounced as the complexity of the intraband system is increased. That intraband SBVQ is also superior from a

Fig. 10. Lenna image coded by previsualized intraband SBVQ with $R = \frac{1}{4}$ and $Q = 2^{16R}\alpha$.

TABLE I
VECTOR DIMENSION AND CODEBOOK SIZE ASSIGNMENTS FOR PREVISUALIZED IMAGE CODING $Q = 2^{16R}\alpha$. ALL SUBBANDS NOT TABULATED WERE ALLOCATED ZERO BITS.

| Subband | Case 1 | Case 2 | Case 3 | Case 4 | Case 5 | Case 6 |
|---|---|---|---|---|---|---|
| 00-00 | k = 4 | k = 4 | k = 4 | k = 4 | k = 4 | k = 4 |
| | N = 212 | N = 424 | N = 817 | N = 1526 | N = 3683 | N = 5973 |
| 00-01, 00-10 | k = 4 | k = 4 | k = 4 | k = 4 | k = 4 | k = 4 |
| | N = 14 | N = 27 | N = 56 | N = 124 | N = 120 | N = 175 |
| 00-11 | k = — | k = — | k = — | k = — | k = — | k = 4 |
| | N = — | N = — | N = — | N = — | N = — | N = 12 |
| 10-10, 01-01 | k = — | k = 16 | k = 16 | k = 16 | k = 8 | k = 16 |
| | N = — | N = 11 | N = 41 | N = 132 | N = 80 | N = 923 |
| Rate | 0.25 | 0.3125 | 0.375 | 0.4375 | 0.5 | 0.5625 |
| Q (Design) | 16 | 32 | 64 | 128 | 256 | 512 |
| Q (Actual) | 15.00 | 31.25 | 63.19 | 127.38 | 255.19 | 511.31 |

TABLE II
VECTOR DIMENSION AND CODEBOOK SIZE ASSIGNMENTS FOR PREVISUALIZED IMAGE CODING $Q = 512\alpha$. ALL SUBBANDS NOT TABULATED WERE ALLOCATED ZERO BITS.

| Subband | Case 1 | Case 2 | Case 3 | Case 4 | Case 5 | Case 6 |
|---|---|---|---|---|---|---|
| 00-00 | k = 4 | k = 4 | k = 4 | k = 4 | k = 4 | k = 4 |
| | N = 1212 | N = 2296 | N = 3760 | N = 6160 | N = 6063 | N = 5973 |
| 00-01, 00-10 | k = 4 | k = 4 | k = 4 | k = 4 | k = 4 | k = 4 |
| | N = 7 | N = 16 | N = 29 | N = 54 | N = 153 | N = 175 |
| 00-11 | k = — | k = — | k = — | k = — | k = — | k = 4 |
| | N = — | N = — | N = — | N = — | N = — | N = 12 |
| 10-10, 01-01 | k = — | k = 16 | k = 16 | k = 16 | k = 16 | k = 16 |
| | N = — | N = 3 | N = 27 | N = 226 | N = 905 | N = 923 |
| Rate | 0.25 | 0.3125 | 0.375 | 0.4375 | 0.5 | 0.5625 |
| Q (Design) | 512 | 512 | 512 | 512 | 512 | 512 |
| Q (Actual) | 76.63 | 145.88 | 242.00 | 420.00 | 511.19 | 511.31 |

TABLE III
VECTOR DIMENSION AND CODEBOOK SIZE ASSIGNMENTS FOR $Q = 512\alpha$ (MVM NOT INCLUDED). ALL SUBBANDS NOT TABULATED WERE ALLOCATED ZERO BITS.

| Subband | Case 1 | Case 2 | Case 3 | Case 4 | Case 5 | Case 6 |
|---|---|---|---|---|---|---|
| 00-00 | k = 4 | k = 4 | k = 4 | k = 4 | k = 4 | k = 4 |
| | N = 2377 | N = 4883 | N = 8060 | N = 7900 | N = 6991 | N = 6824 |
| 00-01, 00-10 | k = 4 | k = 4 | k = 4 | k = 4 | k = 4 | k = 4 |
| | N = 5 | N = 15 | N = 29 | N = 62 | N = 173 | N = 208 |
| 00-11 | k = — | k = — | k = — | k = — | k = — | k = 4 |
| | N = — | N = — | N = — | N = — | N = — | N = 11 |
| 10-10, 01-01 | k = — | k = — | k = 16 | k = 16 | k = 16 | k = 16 |
| | N = — | N = — | N = 6 | N = 79 | N = 422 | N = 466 |
| Rate | 0.25 | 0.3125 | 0.375 | 0.4375 | 0.5 | 0.5625 |
| Q (Design) | 512 | 512 | 512 | 512 | 512 | 512 |
| Q (Actual) | 149.19 | 307.06 | 508.13 | 511.38 | 511.31 | 511.44 |

Fig. 11. Lenna image coded by previsualized intraband SBVQ with $R = \frac{1}{4}$ and $Q = 512\alpha$.

Fig. 12. Lenna image coded by intraband SBVQ with $R = \frac{9}{16}$ and $Q = 512\alpha$.

subjective viewpoint is seen by comparing Figures 7–11.

The experiments depicted in Table 2 were repeated without including the MVM, i.e., with the intraband system designed to minimize MSE directly. The resulting vector dimension and codebook size assignments are shown in Table 3, which demonstrates how the usage of the MVM resulted in a greater allocation of resources to the higher frequency subbands. The compressed image for the case of $R = \frac{9}{16} bpp$ is shown in Figure 12; this image has a PSNR of 30.2 dB. The advantage of using the MVM is apparent by comparing this result with Figure 9.

## VI. Conclusion

This work investigated the utilization of VQ in SBC. The major contribution is the development of a design algorithm for the intraband coder that allocates available resources among the various vector quantizers in an optimal fashion. The algorithm will select a vector dimension and a codebook size for each subband such that constraints on overall bit-rate and encoding complexity are satisfied. The experimental results indicated that the optimal intraband SBVQ coder yields higher quality quantized images (subjectively as well as quantitatively) than interband SBVQ; the improvement becomes more pronounced as the computational complexity of the intraband encoder is allowed to increase. Experimental results were also included that demonstrated the effect of the MVM on intraband system performance.

An effect present in the intraband SBVQ quantized images is the "ringing" that emanates from the edges. This is due to quantization error in the image subbands that in turn results in aliasing error during the subband synthesis process. The recent work of [31] investigated this phenomenon in conjunction with PCM coding of subband signals where it was concluded that spatially adaptive quantization would reduce this distortion. Further work is needed to mitigate the problem for intraband SBVQ.

The visual model utilized in this work was a single channel system describing the actions of the retina. It should be noted, however, that electrophysiological experiments on the cerebral cortex have revealed the presence of multichannel processing in the higher stages of the visual system [32]: an observed image is separated into a number of neural signals that represent different image frequency components. Although the subband decomposition was used here as a means of reducing MSE within the visual space defined by the MVM, the multichannel nature of a subband coder can be further utilized in exploiting perceptual phenomena that occur as a consequence of the higher level processing. Reference [33] describes an extension of this work in which the bit allocation technique presented here is used to design an intraband SBVQ coder that incorporates visual masking functions.

## References

[1] R. E. Crochiere, S. A. Webber, and J. L. Flanagan, "Digital coding of speech in sub-bands," *Bell System Tech. J.*, pp. 1069–1085, Oct. 1976.

[2] D. Esteban and C. Galand, "32 kb/s CCITT-compatible split band coding scheme," in *Proc. Int'l. Conf. Acoust., Speech, Signal Process.*, pp. 320–325, IEEE Signal Process. Soc., 1978.

[3] Y. Huang and P. M. Schultheiss, "Block quantization of correlated gaussian random variables," *IEEE Trans. Comm. Sys.*, pp. 289–296, Sept. 1976.

[4] N. S. Jayant and P. Noll, *Digital Coding of Waveforms.* Englewood Cliffs, NJ: Prentice Hall, 1984.

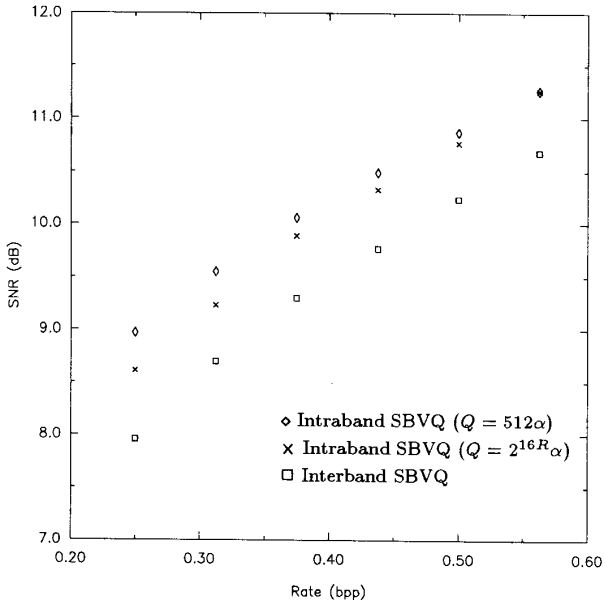


Fig. 13. Previsualized image SBVQ rate-distortion characteristics.

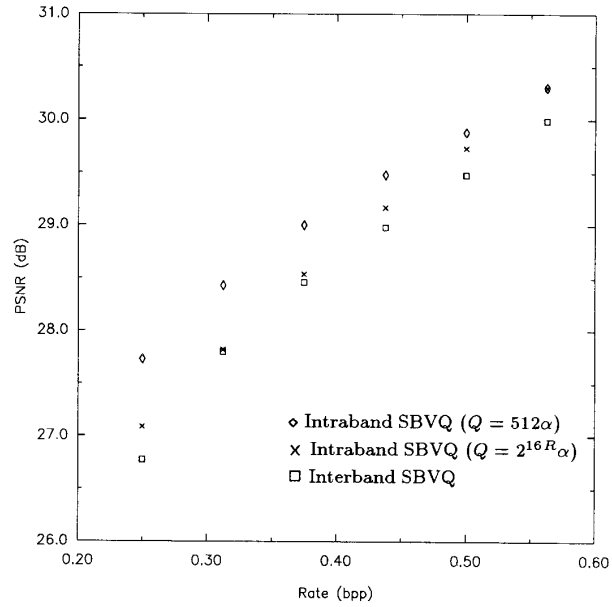


Fig. 14. SBVQ rate-distortion characteristics.


[5] R. M. Gray, "Vector quantization," *IEEE Acoust., Speech, Signal Process. Mag.*, vol. 1, pp. 4–29, Apr. 1984.

[6] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. 28, pp. 84–95, Jan. 1980.

[7] J. Makhoul, S. Roucos, and H. Gish, "Vector quantization in speech coding," *Proc. IEEE*, vol. 73, pp. 1551–1588, Nov. 1985.

[8] J. N. Bradley, *Subband image coding with vector quantization.* PhD thesis, Dept. of Electrical Engineering, University of Utah, Salt Lake City, UT, Dec. 1989.

[9] V. Cuperman, "On adaptive vector transform quantization for speech coding," *IEEE Trans. Commun.*, vol. 37, pp. 261–267, Mar. 1989.

[10] P. H. Westerink, D. E. Boekee, J. Biemond, and J. W. Woods, "Subband coding of images using vector quantization," *IEEE Trans. Commun.*, vol. 36, pp. 713–719, June 1988.

[11] M. Rabbani and P. W. Jones, *Digital Image Compression Techniques.* No. 7 in SPIE Tutorial Texts in Optical Engineering, Bellingham, WA: Soc. Photo-Opt. Instrument. Engin., 1991.

[12] H. Abut and S. A. Luse, "Vector quantizers for subband coded waveforms," in *Proc. Int'l. Conf. Acoust., Speech, Signal Process.*, Apr. 1984. paper 10.6.

[13] H. Abut and S. Ergezinger, "Low-rate encoding using vector quantization and subband coding," in *Proc. Int'l. Conf. Acoust., Speech, Signal Process.* [34], pp. 449–452.

[14] P. Chang, R. M. Gray, and J. May, "Fourier transform vector quantization for speech coding," *IEEE Trans. Comm.*, pp. 1059–1068, Oct. 1987.

[15] A. Gersho, T. Ramstad, and I. Versvik, "Fully vector-quantized subband coding with adaptive codebook allocation," in *Proc. Int'l. Conf. Acoust., Speech, Signal Process.*, Apr. 1984. paper 10.7.

[16] I. Versvik and H. C. Guren, "Suband coding with vector quantization," in *Proc. Int'l. Conf. Acoust., Speech, Signal Process.* [34], pp. 3099–3102.

[17] P. H. Westerink, J. Biemond, and D. E. Boekee, "An optimal bit allocation algorithm for sub-band coding," in *Proc. Int'l. Conf. Acoust., Speech, Signal Process.*, (New York, NY), IEEE Signal Process. Soc., Apr. 1988.

[18] N. Tanabe and N. Farvardin, "Subband image coding using entropy-coded quantization over noisy channels," in *Proc. Int'l. Conf. Acoust., Speech, Signal Process.*, (Albuquerque, NM), pp. 2105–2108, IEEE Signal Process. Soc., Apr. 1990. Paper M6.7.

[19] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing.* Englewood Cliffs, NJ: Prentice Hall, 1983.

[20] J. W. Woods and S. D. O'Neil, "Subband coding of images," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, pp. 1278–1288, Oct. 1986.

[21] W. H. Equitz, "A new vector quantization clustering algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, pp. 1568–1575, Oct. 1989.

[22] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Info. Theory*, vol. 25, pp. 373–380, July 1979.

[23] A. Segall, "Bit allocation and encoding for vector sources," *IEEE Trans. Info. Theory*, vol. 22, pp. 162–169, Mar. 1976.

[24] P. E. Gill, W. Murray, M. A. Saunders, and M. H. Wright, "User's guide for NPSOL (version 4.0): a Fortran package for nonlinear programming," Tech. Rep. SOL 86-2, Systems Optimization Lab, Dept. of Operations Research, Stanford Univ., Jan. 1986.

[25] T. G. Stockham, Jr., "Intra-frame encoding for monochrome images by means of a psychophysical model based on nonlinear filtering of multiplied signals," in *Proc. 1969 Symp. Picture Bandwidth Compression* (T. S. Huang and O. J. Tretiak, eds.), (New York, NY), Gordon and Breach, 1972.

[26] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing.* Englewood Cliffs, NJ: Prentice Hall, 1989.

[27] T. G. Stockham, Jr., "The application of generalized linearity to automatic gain control," *IEEE Trans. Audio and Electroacoustics*, June 1968.

[28] T. G. Stockham, Jr., "Image processing in the context of a visual model," *Proc. IEEE*, vol. 60, pp. 828–842, July 1972.

[29] Z. Xie and T. G. Stockham, Jr., "Toward the unification of three visual laws and two visual models in brightness perception," *IEEE Trans. Systems, Man, Cybernet.*, vol. 19, pp. 379–387, March/April 1989.

[30] P. Colas-Baudelaire, *Digital picture processing and psychophysics.* PhD thesis, University of Utah, 1973.

[31] P. H. Westerink, J. Biemond, and D. E. Boekee, "Scalar quantization error analysis for image subband coding using qmf's," *ieee-tsp*, pp. 421–428, Feb. 1992.

[32] D. H. Hubel and T. N. Wiesel, "Brain mechanisms of vision," *Scientific American*, pp. 150–162, Sept. 1979.

[33] R. Baseri and V. J. Mathews, "Vector quantization of images using visual masking functions," in *Proc. Int'l. Conf. Acoust., Speech, Signal Process.*, (San Francisco, CA), pp. III–365–III–368, IEEE Signal Process. Soc., Mar. 1992.

[34] *Proc. Int'l. Conf. Acoust., Speech, Signal Process.*, (Tokyo, Japan), IEEE Signal Process. Soc., Apr. 1986.

**Jonathan N. Bradley** (S'80-'82) was born in Washington, D.C. on November 5, 1959. He received the B.S. and M.S. degrees in Electrical Engineering from the University of Wyoming in 1980 and 1982, respectively. He received the Ph.D. degree in Electrical Engineering from the University of Utah in 1989.

From 1982-1984 he was a Technical Staff Member at Hughes Aircraft Company Missile Systems Group, Tucson, Arizona. From 1989-1990 he was a postdoctoral fellow in the Department of Electrical and Computer Engineering at the University of Victoria, Victoria, B.C., Canada. Since 1990 he has been employed in the Computer Research and Applications Group at Los Alamos National Laboratory, Los Alamos, New Mexico.

**Thomas G. Stockham, Jr.** (S'55-M'60-SM'69-F'77) was born in Passaic, N.J. on December 22, 1933. He received the S.B., S.M., and Sc.D. degrees in 1955, 1956, and 1959, respectively, from the Massachusetts Institute of Technology, Cambridge, MA.

From 1955 to 1959, concurrent with his graduate studies, he was a Teaching Assistant, receiving the Goodwin Teaching Award in 1957. In 1959, he was appointed Assistant Professor in the Department of Electrical Engineering, Massachusetts Institute of Technology. Since 1958, he has centered his research in the field of information processing. From 1966 to 1968, he was a Staff Member of the M.I.T. Lincoln Laboratory, Lexington, MA, where he developed nonlinear systems for image enhancement and audio compression using the principle of generalized superposition. In July 1968, he was appointed to the Computer Science faculty at the University of Utah, Salt Lake City, UT, where he and his students developed methods for modeling human vision; graphics image shading; image deblurring and compression; sound dereverberation; and color image enhancement. In 1973-1974 he served on a panel of experts for Chief Judge John J. Sirica to examine the Watergate Tapes.

In 1975 Dr. Stockham founded Soundstream, Inc. Under his direction the company developed digital commercial sound recording and editing. These developments led to the establishment of these practices industry wide. The resulting accumulation of more than 100 digitally recorded albums by early 1981 triggered the advent of the compact disc (CD).

In 1983 he rejoined the University of Utah faculty in the Department of Electrical Engineering where he has since been developing improved visual models and applying them to image compression methods using vector quantization. He is also an active consultant in digital image and sound processing technologies.

Dr. Stockham is a member of Tau Beta Pi, Sigma Xi, Eta Kappa Nu, and the Association for Computing Machinery. He is a Fellow of the IEEE and the Audio Engineering Society and served as president of the latter in 1982-1983. He received the IEEE Group on Audio and Electroacoustics Senior Award in 1969, the IEEE Utah Chapter Award in 1972, 1978, and 1985, the SMPTE Alexander M. Poniatoff Gold Medal for Technical Excellence in 1985, the Utah Engineers Council Engineer of the Year Award and the University of Utah College of Engineering Outstanding Teacher Award in 1986, the Audio Engineering Society Gold Medal in 1987, and a National Academy of Television Arts and Sciences "Emmy" Award in 1988.

**V. John Mathews** (S'82-M'84-SM'90) received his B.E. (hons.) degree in Electronics and Communication Engineering from the University of Madras, India, and the M.S., and Ph.D. degrees in Electrical and Computer Engineering from the University of Iowa, Iowa City, in 1980, 1982 and 1984, respectively.

From 1980 to 1984 he held a Teaching-Research Fellowship at the University of Iowa, where he also worked as a Visiting Assistant Professor with the Department of Electrical and Computer Engineering from 1984 to 1985. He joined the Department of Electrical Engineering at the University of Utah, Salt Lake City in 1985 and is currently an Associate Professor there. His research interests include adaptive filtering, data compression and spectrum estimation.

Dr. Mathews is a Senior Member of the IEEE. He was an Associate Editor of the *IEEE Transactions on Signal Processing* during 1989-91 and is an Associate Editor of *IEEE Signal Processing Letters* now. He serves on the Digital Signal Processing Technical Committee of the IEEE Signal Processing Society.