Subband Coding of Images

JOHN W. WOODS, SENIOR MEMBER, IEEE, AND SEAN D. O'NEIL

Abstract—**Subband coding has become quite popular for the source encoding of speech. This paper presents a simple yet efficient extension of this concept to the source coding of images. We specify the constraints for a set of two-dimensional quadrature mirror filters (QMF's) for a particular frequency-domain partition, and show that these constraints are satisfied by a separable combination of one-dimensional QMF's. Bits are then optimally allocated among the subbands to minimize the mean-squared error for DPCM coding of the subbands. Also, an adaptive technique is developed to allocate the bits within each subband by means of a local variance mask. Optimum quantization is employed with quantizers matched to the Laplacian distribution. Subband coded images are presented along with their signal-to-noise ratios (SNR's). The SNR performance of the subband coder is compared to that of the adaptive discrete cosine transform (DCT), vector quantization, and differential vector quantization for bit rates of 0.67, 1.0, and 2.0 bits per pixel for 256 × 256 monochrome images. The adaptive subband coder has the best SNR performance.**

I. Introduction

SUBBAND coding of speech was introduced by Crochiere *et al.* [1] in 1976. Since that time, this technique has become quite popular for the medium bandwidth waveform coding of speech. The basic idea of subband coding (SBC) is to split up the frequency band of the signal and then to code each subband with either PCM or DPCM using a coder and bit rate accurately matched to the statistics of that band. In addition to the obvious advantages of such an approach, two other advantages have been found important in speech coding [2]. First, the error in coding a subband is confined to that subband, thus exploiting the masking effect of speech. Second, by varying the bit assignment among the subbands, the noise spectrum can be shaped according to the subjective noise perception of the human ear.

Related previous work in multiband source encoding of images traces back to the splitband quantization method of Kretzmer [3] in 1956. This was a high-frequency–low-frequency PCM-type approach. Another two-band system was the synthetic highs system introduced by Schreiber *et al.* [4] in 1959. More directly related to our method is the recent pyramid coding method of Burt and Adelson [5]. Their approach used a Laplacian pyramid to produce an approximate frequency decomposition but did not explicitly use any subband coding concepts or QMF filters. Independent of the present work [6], the QMF filter was extended to the multidimensional (m-D) case earlier by Vetterli [7] who claimed that separable filters are necessary and sufficient for the most natural four-band extension of the standard two-band QMF filters introduced by Esteban and Galand [8]. Actually, separability is not necessary, and a weaker condition is derived in Section II. It should be noted that no results on image coding are contained in [7].

This paper is organized as follows. Section II summarizes our extension of subband filters of the QMF variety to two dimensions. Section III introduces the concept of subband coding of images. Section IV considers an adaptive extension motivated by previous work in adaptive transform image coding [9]. Section V presents simulation results including a signal-to-noise ratio (SNR) comparison with several other popular encoding methods extending the comparison reported earlier by Baker and Gray [10]. Section VI presents conclusions and suggestions for further work.

II. Subband Filters in Two Dimensions

This section will summarize our extension of the subband filtering concept to the 2-D case including the method of QMF filter design to eliminate possible aliasing error due to nonideal subband filters. An independent extension of QMF's to the multidimensional case has been published by Vetterli [7]. A key difference is that our derivation substitutes a weaker symmetry constraint for separability of the QMF's, thus allowing the possibility of more general frequency-domain decompositions.

Ideal Filters

Fig. 1 shows the initial 4-band partitioning stage that is the basis for the 16-band filter system to be used in the coding of Section III. After each subband has been split off, it is demodulated to baseband (i.e., zero frequency) by a (2, 2) subsampling [Fig. 2(a)], that is, a decimation by the factor 2 in each dimension. This decimation then makes each subband image fullband at the lower sample rate. For a 16-band system, this process is repeated to further split each subband into four more subbands. It can be seen that the region of support of the resulting subband images would then be reduced by a factor of 4 in each dimension, giving an overall factor of 16 reduction in sample rate for each of the 16 subband images.

After encoding, transmission, and decoding, the filtering process is carried out in reverse [Fig. 2(b)]. Each subband is upsampled by (2, 2) and suitably bandpass filtered

Manuscript received June 22, 1985; revised February 28, 1986.

J. W. Woods is with the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180-3590.

S. D. O'Neil was with the Department of Electrical, Computer, and Systems Engineering, Rennselaer Polytechnic Institute, Troy, NY. He is now with the MITRE Corporation, Bedford, MA 07130.

IEEE Log Number 8609292.

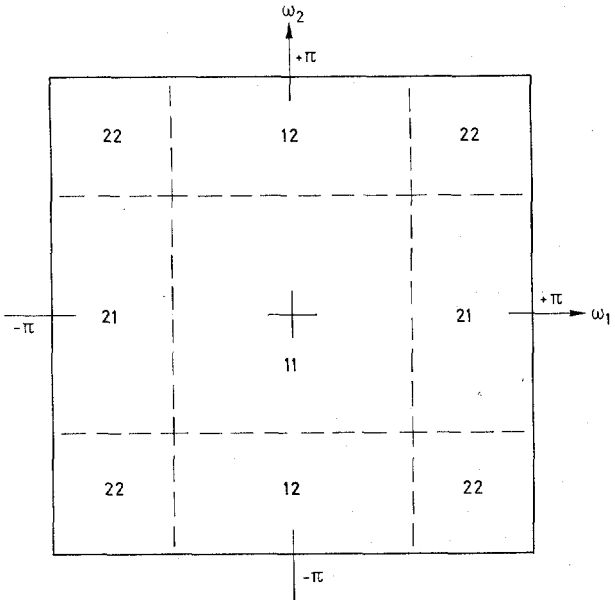


Fig. 1. Initial 4-band partition of frequency domain.

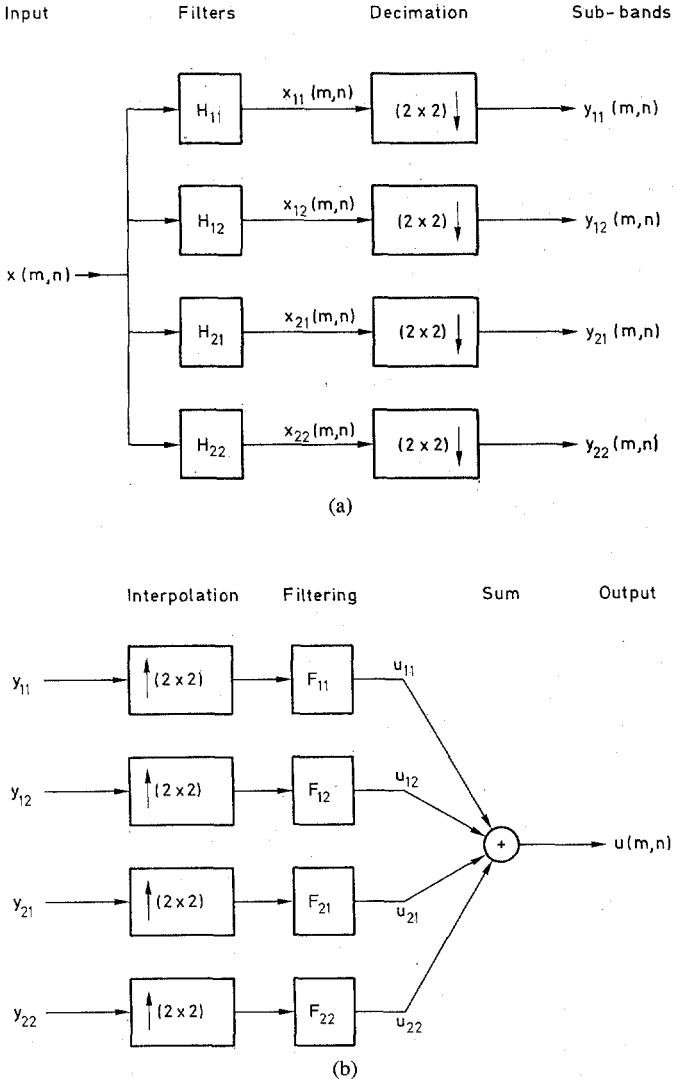


Fig. 2. (a) System diagram of 4-band splitting (done twice for 16 bands). (b) System diagram of 4-band recombination (done twice for 16 bands).

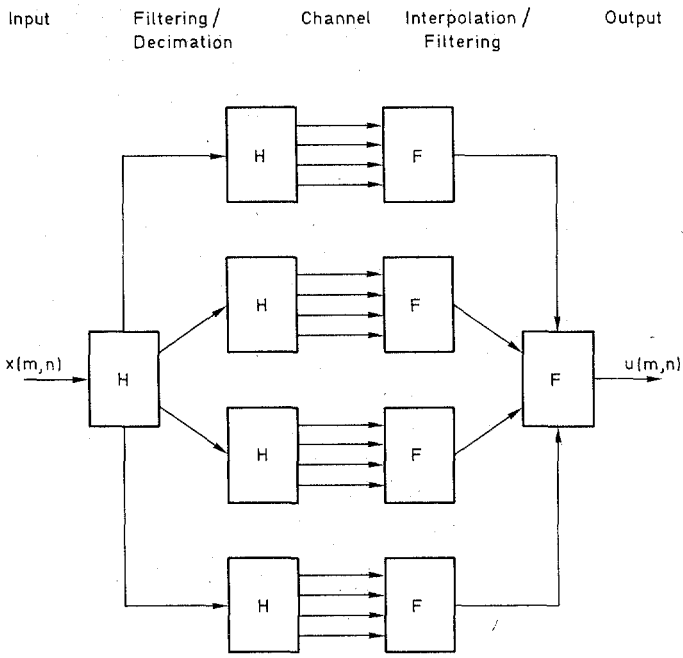


Fig. 3. System diagram of full 16-band system.

to eliminate aliased copies of the signal produced by the upsampling. The appropriate sets of subbands are then summed to reconstruct the original signal. Again, for 16 bands, this process must be repeated once. A schematic of the full system can be seen in Fig. 3.

FIR Filters

When FIR filters are used to approximate the ideal subband filter characteristics of Fig. 1, either gaps or aliasing error will occur due to the effect of decimation in the transition band of the filter. This effect is well known in the 1-D subband filtering literature [2] where the audio effect of either degradation is deemed unacceptable. We have found these distortions to be similarly unacceptable for subband filtering of images as seen in Fig. 4. Here linear taper filters (i.e., a straight line transition band) are used to partition the subbands and ideal filters are used to recombine them. In the 1-D case, the QMF approach was introduced to cancel out the aliasing effect of the filter transition band in the desirable case where no frequency gaps are allowed between the bands. The extension of the QMF concept to two dimensions is relatively straightforward and proceeds as follows.

Referring to Fig. 2(a) and the initial four-band splitting, we require that the four subband filters H_{11} through H_{22} have mirror-image conjugate symmetry about their mutual boundaries, which for real h_{ij} is equivalent to

$$H_{12}(\omega_1, \omega_2) = H_{11}(\omega_1, \omega_2 + \pi) \tag{1a}$$

$$H_{21}(\omega_1, \omega_2) = H_{11}(\omega_1 + \pi, \omega_2) \tag{1b}$$

$$H_{22}(\omega_1, \omega_2) = H_{11}(\omega_1 + \pi, \omega_2 + \pi). \tag{1c}$$

Denoting the outputs of the filters in Fig. 2(a) as x_{11} through x_{22}, we have the outputs Y_{ij} after (2, 2) downsampling,

Fig. 4. Reconstruction of "Lady" image using linear taper filters for splitting and ideal filters for reconstruction (convolution performed in DFT frequency domain).

$$Y_{ij}(\omega_1, \omega_2) = \frac{1}{4} \sum_{k=0}^{1} \sum_{l=0}^{1} H_{ij}\left(\frac{\omega_1 + k\pi}{2}, \frac{\omega_2 + l\pi}{2}\right) \cdot X\left(\frac{\omega_1 + k\pi}{2}, \frac{\omega_2 + l\pi}{2}\right). \quad (2)$$

Ignoring the effects of encoding/decoding and transmission, we find that the outputs U_{ij} of the interpolation filters [Fig. 2(b)] can be expressed as

$$U_{ij}(\omega_1, \omega_2) = Y_{ij}(2\omega_1, 2\omega_2)\, F_{ij}(\omega_1, \omega_2),$$

and that the final output U can be written as

$$U(\omega_1, \omega_2) = \tfrac{1}{4} \sum_{k,l} X(\omega_1 + k\pi, \omega_2 + l\pi) \cdot \left[\sum_{i,j} H_{ij}(\omega_1 + k\pi, \omega_2 + l\pi)\, F_{ij}(\omega_1, \omega_2)\right], \quad (3)$$

which is essentially the same as (3.2) in [7]. Splitting U into a desired signal component U_o and an undesired aliasing component U_a such that $U = U_o + U_a$, we get the aliased component

$$U_a(\omega_1, \omega_2) = \tfrac{1}{4} \sum_{(k,l) \neq (0,0)} X(\omega_1 + k\pi, \omega_2 + l\pi) \cdot \left[\sum_{i,j} H_{ij}(\omega_1 + k\pi, \omega_2 + l\pi)\, F_{ij}(\omega_1, \omega_2)\right],$$

which will vanish if and only if (iff)

$$\sum_{i,j} H_{ij}(\omega_1 + k\pi, \omega_2 + l\pi)\, F_{ij}(\omega_1, \omega_2) = 0 \quad \text{for } (k, l) \neq (0, 0). \quad (4)$$

We chose the same reconstruction filters F_{ij} as Vetterli [7],

$$F_{11}(\omega_1, \omega_2) = 4H_{11}(\omega_1, \omega_2) \quad (5a)$$

$$F_{12}(\omega_1, \omega_2) = -4H_{12}(\omega_1, \omega_2) \quad (5b)$$

$$F_{21}(\omega_1, \omega_2) = -4H_{21}(\omega_1, \omega_2) \quad (5c)$$

$$F_{22}(\omega_1, \omega_2) = 4H_{22}(\omega_1, \omega_2), \quad (5d)$$

and found that the aliased terms automatically vanish for $(k, l) = (0, 1)$ or $(1, 0)$. However, in the case $(k, l) = (1, 1)$ the aliased term will vanish iff

$$H_{11}(\omega_1, \omega_2)\, H_{11}(\omega_1 + \pi, \omega_2 + \pi) = H_{11}(\omega_1, \omega_2 + \pi)\, H_{11}(\omega_1 + \pi, \omega_2). \quad (6)$$

Since the four factors in (6) are evaluations of H_{11} on a square of side π, it follows that both sides of (6) will be zero if the support of H_{11} is contained within $(-\frac{1}{2}\pi, +\frac{1}{2}\pi) \times (-\frac{1}{2}\pi, +\frac{1}{2}\pi)$ whether H_{11} is separable or not, thus contradicting the claim for the necessity of separability in [7]. Note, however, that any separable H_{11} would clearly satisfy (6). An advantage of nonseparable filters would be a directional capability in the subband decomposition. The advantage of separability is, of course, computational simplicity.

Employing a linear phase symmetric $L \times L$ FIR filter for h_{11} with L *even*, we have

$$h_{11}(m, n) = h_{11}(L - 1 - m, L - 1 - n), \quad 1 \leq m, \quad n \leq \tfrac{1}{2}L - 1,$$

so that, on removing the linear phase factor, the nonaliased term becomes

$$|H_{11}^2(\omega_1, \omega_2)| + |H_{11}^2(\omega_1, \omega_2 + \pi)| + |H_{11}^2(\omega_1 + \pi, \omega_2)| + |H_{11}^2(\omega_1 + \pi, \omega_2 + \pi)| = 1. \quad (7)$$

Thus, a general 2-D QMF would be generated by a lowpass filter optimized to approximately satisfy (7) with the constraint of (6).

Avoiding this potentially difficult design problem, we proceed here in this first investigation of subband image coding, with the easier separable case. A 1-D QMF filter pair h_1 and h_2 satisfies [8]

$$h_1(n) = h_1(L - 1 - n) \quad 0 \leq n \leq L/2 - 1 \quad (8a)$$

$$h_2(n) = (-1)^n h_1(n) \quad (8b)$$

$$|H_1^2(\omega)| + |H_2^2(\omega)| = 1. \quad (8c)$$

If we define the 2-D baseband filter h_{11} as the separable product

$$h_{11}(m, n) \triangleq h_1(m)\, h_1(n), \quad (9)$$

take the Fourier transform, and substitute into (6), we see that the 2-D QMF filters can be taken as a separable product of identical 1-D QMF filters [7]

$$H_{ij}(\omega_1, \omega_2) = H_i(\omega_1)\, H_j(\omega_2), \quad 1 \leq i, j \leq 2. \quad (10)$$

In our coding simulations, we used the 32-point QMF

Fig. 5. Reconstruction of "Lady" image using 16-band 2-D QMF filter (convolution performed in DFT frequency domain).

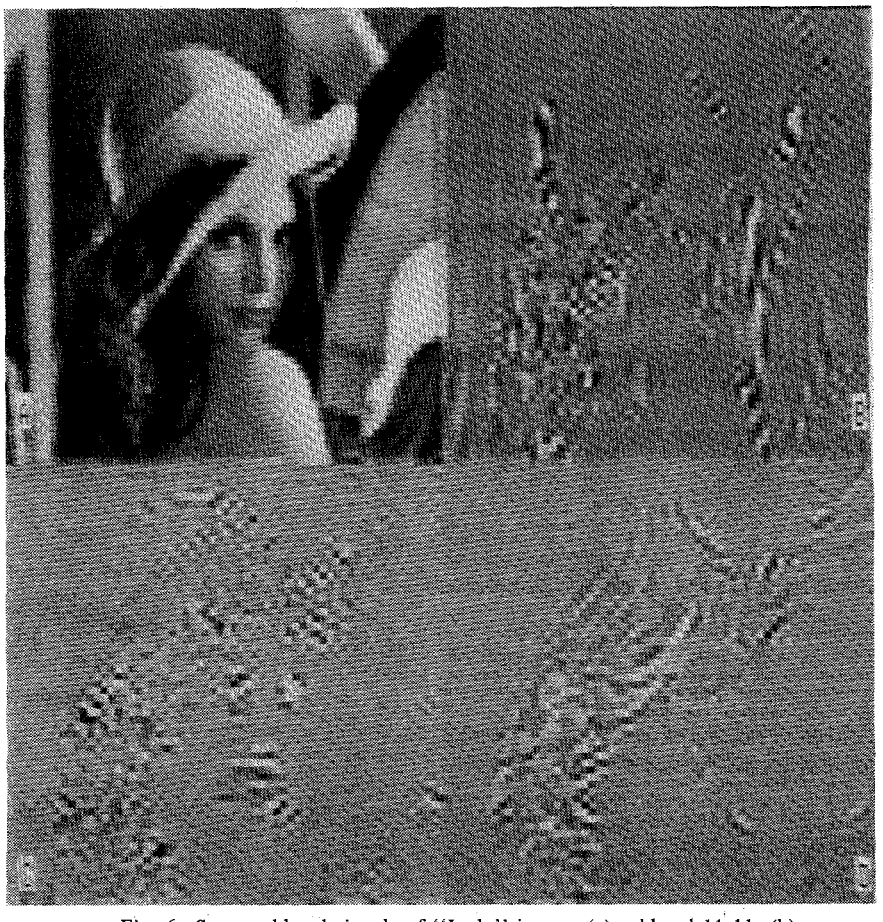


Fig. 6. Some subband signals of "Lady" image: (a) subband 11-11, (b) subband 11-21, (c) subband 11-12, (d) subband 11-22 (zoomed by 2×2).

designated as $32D$ in [11]. It has a transition bandwidth of 0.043 rad and an overall passband ripple of 0.025 dB. The stopband rejection varies from 38 to 48 dB. Fig. 5 shows a 16-band reconstruction obtained using the separable 2-D QMF based on this 1-D filter. The filtering was performed in the DFT frequency domain with $N = 256$, the image size. Comparison to the original shows it to be virtually identical. Fig. 6 shows four of the subband im-

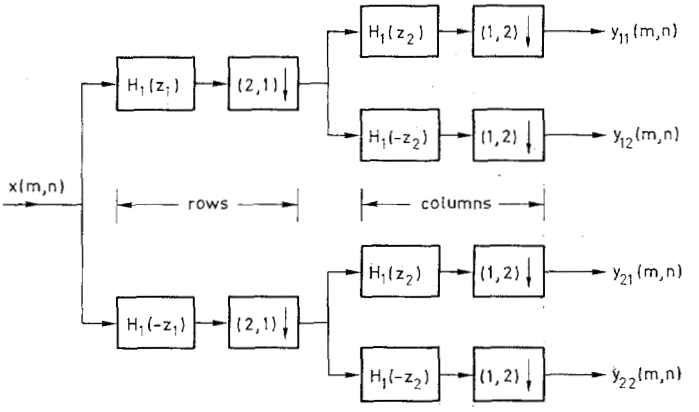


Fig. 7. Separable 4-subband filter.

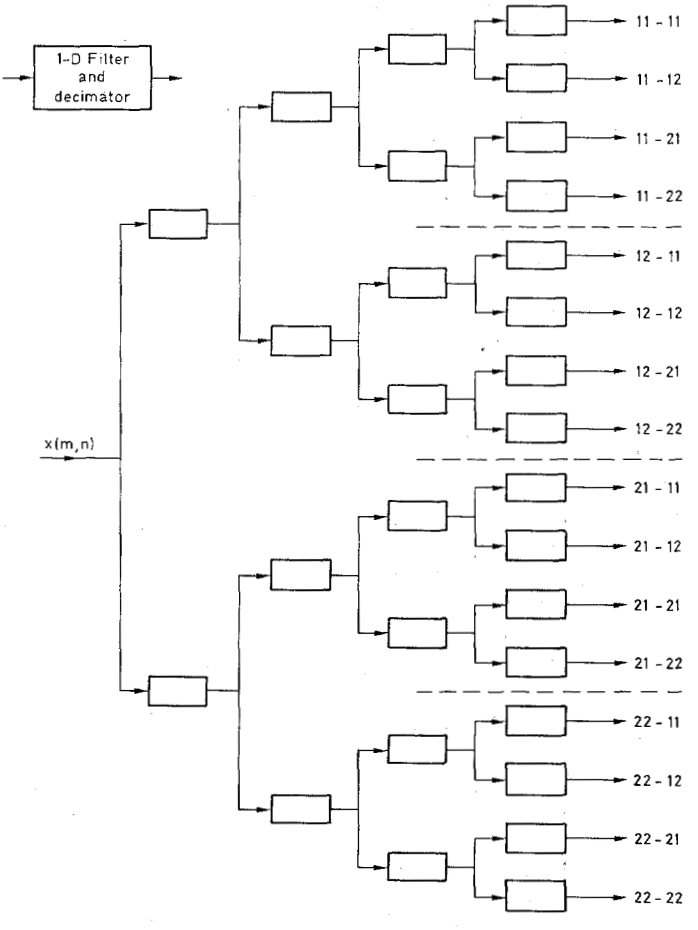


Fig. 8. Tree structure for 16-band separable subband filter.

ages in this decomposition. Fig. 6(a) shows the baseband image 11-11. Fig. 6(b) and (c) shows subband images 11-12, -21, and -22 which have been scaled up for presentation. The rapid black–white transitions in these subband images result from the demodulation by decimation. The frequency axis is effectively folded about the lower band edge, thus interchanging the high and low frequencies. For these subbands, this is equivalent to multiplying by $(-1)^m$, $(-1)^n$, and $(-1)^{m+n}$, respectively. This effect, although visually annoying, would cause no problem for a DPCM predictor.

Filter Tree

Using the 1-D QMF filters, we can perform the 4-band decomposition by the row–column approach shown in Fig. 7. This can be regarded as the first two levels in a filter tree structure with the understanding that the first stage represents row filtering and the second stage represents column filtering. To get a 16-band decomposition, as mentioned above, we iterate the four-band decomposition once to get the filter tree shown in Fig. 8. Here each box represents a 1-D filtering with $h_1(n)$ or $(-1)^n h_1(n)$ followed by a $2\downarrow$ subsampling. Also in this diagram, we have

the convention that the odd stages represent row filtering while the even stages represent column filtering.

To calculate the computational requirements of this filter tree, we note that stage 1 will require $\frac{1}{2} LN^2$ multiplies where $N \times N$ is the image size and L is the filter length which is assumed to be even. At the second stage, we have one-half as many columns which must be filtered in the same way, but there are two branches, thus, we again require $\frac{1}{2} LN^2$ multiplies. Proceeding to the last two stages we obtain the same results. Thus, the total multiplication requirement becomes $2LN^2$ where we have assumed that the multiplications for filtering with $h_1(n)$ and $(-1)^n h_1(n)$ are performed just once and then added or subtracted as appropriate into the relevant partial product. For $L = 32$, this is 64 multiplies/pixel. These filter trees were also considered by Vetterli [7], who using the concept of pseudo-QMF has reported 10 multiplies/pixel for this four stage filter with $L = 25$. For comparison, a 16×16 DCT using the method of [9] requires about $5\frac{1}{2}$ multiplies/pixel.

III. Subband Coding of Images

This section treats the encoding of the subbands created by the QMF filter tree of the last section. We will be mainly concerned with DPCM as the method to encode the subbands rather than PCM which is often used in subband speech encoders. Our use of DPCM is motivated by the increased efficiency of a predictive encoder for a nonwhite power spectral density. The overall subband coding system is shown in Fig. 9.

In the case of M subbands of equal bandwidth, each subband has been subsampled by sqrt(M) in each dimension. Thus, if we assign B_k bits to subband k, we have the following average bit rate:

$$B = \frac{1}{M} \sum_{k=1}^{M} B_k. \tag{11}$$

Here k represents the subbands ij of the previous section indexed in some convenient order.

To design the DPCM coder for each subband, we need a linear predictive or AR model in the form

$$y_k(m, n) = \sum_{o,p} c_{op}^k \, y_k(m - o, n - p) + w_k(m, n)$$

with prediction error variance $\sigma_{p,k}^2$. Referring to Fig. 10, we see that for DPCM the reconstruction mean-square error (MSE) is equal to the quantizer MSE

$$\begin{aligned} E[(y_k(m, n) - \tilde{y}_k(m, n))^2] &= E[(e_k(m, n) - q_k(m, n))^2] \\ &= \text{Var}\,[r_k(m, n)] \triangleq \sigma_{r,k}^2 \end{aligned}$$

where $r_k \triangleq y_k - \tilde{y}_k$ is the reconstruction error.

From [2] this is given by

$$\begin{aligned} \sigma_{r,k}^2 &= g(B_k)\ \text{Var}\,[e_k(m, n)]\ 2^{-2B_k} \\ &\simeq g(B_k)\ \sigma_{p,k}^2\, 2^{-2B_k}, \end{aligned} \tag{12}$$

where $g(B_k)$ is a slowly varying function of the number of bits assigned to subband k. The histogram of prediction errors in subband coding (SBC) of images matches the

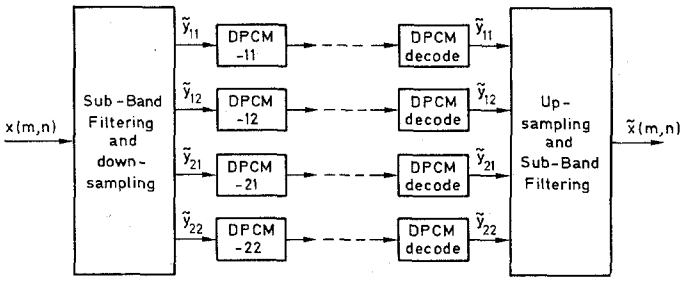


Fig. 9. A subband coding system with four bands.

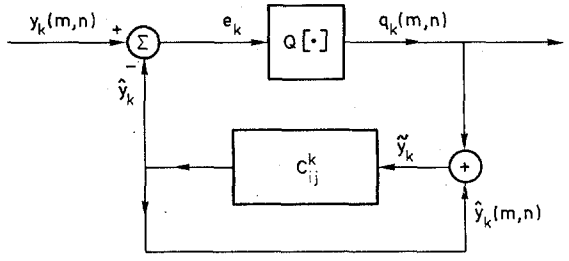


Fig. 10. DPCM for subband k.

Laplacian probability density function (pdf) quite well (see Fig. 11), a property that it no-doubt inherits from fullband DPCM. The optimal pdf optimized quantizer for the Laplacian density was derived by Nitadori [12], and we employ his results in this paper. The value $g(B_k) \simeq g = 4.5$ is valid for this pdf. We can thus write the sum of the error variances of the M subbands as

$$\sum_{k=1}^{M} \text{Var}\,[r_k(m, n)] = \sum_{k=1}^{M} g\, \sigma_{p,k}^2\, 2^{-2B_k}. \tag{13}$$

This MSE is related to the total reconstruction MSE as

$$\begin{aligned} \text{Var}\,(r) &= E[(y(m, n) - \tilde{y}(m, n))^2] \\ &= \sum_{k=1}^{M} \text{Var}\,(r_k). \end{aligned} \tag{14}$$

Thus, we must minimize (13) subject to the constraint of (11). This is a standard problem in the optimization of transform encoders [9] whose solution is obtained using Lagrange multipliers. The MSE optimal bit assignment then is

$$B_k = B + \frac{1}{2} \log_2 \left[\frac{\sigma_{p,k}^2}{\sigma_{gm}^2} \right], \quad 1 \le k \le M \tag{15}$$

where $\sigma_{gm}^2 \triangleq (\Pi_{k=1}^{M} \sigma_{p,k}^2)^{1/M}$ is the geometric mean of the $\sigma_{p,k}^2$. The minimum overall MSE, then becomes

$$\text{Var}_o(r) = M\, \sigma_{g,m}^2\, 2^{-2B}.$$

Calculation of Coding Gain

The coding gain of fullband DPCM with respect to PCM is defined as the ratio of the PCM reconstruction MSE to that of DPCM [2]. Denoting this quantity by $G_{\text{DPCM/PCM}}$, we can express it as

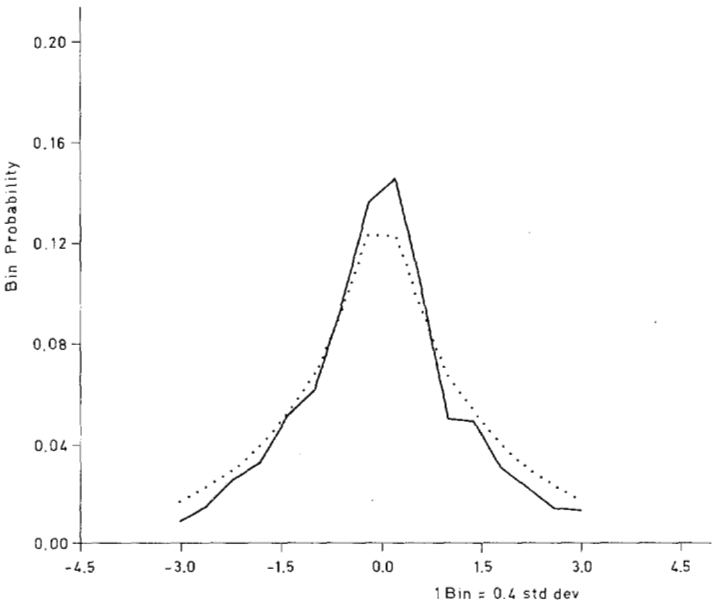


Fig. 11. Histogram of prediction error compared to Laplacian pdf.

$$G_{\text{DPCM/PCM}} = \frac{g\,\sigma_y^2\,2^{-2B}}{g\,\sigma_p^2\,2^{-2B}} = \frac{\sigma_y^2}{\sigma_p^2} \quad \text{(for Gaussian case only)}, = \gamma_y^{-2},$$

where γ_y^2 is the spectral flatness measure [2] given by

$$\gamma_y^2 \triangleq \frac{\exp\frac{1}{2\pi}\left(\int_{-\pi}^{+\pi} \ln S_y(\omega)\,d\omega\right)}{\sigma_y^2},$$

and where σ_p^2 is the fullband DPCM prediction error. The 2-D version of this result follows from the analogous result that the 2-D prediction error can be expressed as

$$\sigma_p^2 = \exp\left(\frac{1}{(2\pi)^2}\int_{-\pi}^{+\pi}\int_{-\pi}^{+\pi} \ln S_y(\omega_1, \omega_2)\,d\omega_1\,d\omega_2\right) \tag{16}$$

If we consider the coding gain of SBC using DPCM to encode the subbands versus fullband DPCM, we have

$$G_{\text{SBC/DPCM}} = \frac{g\,\sigma_p^2\,2^{-2B}}{g\sum_{k=1}^{M}\sigma_{p,k}^2\,2^{-2B}}.$$ [1]

Evaluating at the optimal bit assignment (15) we get

$$G_{\text{SBC/DPCM}} = \frac{\sigma_p^2}{M\sigma_{g,m}^2} \tag{17}$$

where $\sigma_{g,m}^2$ was defined above. Surprisingly, this quantity is exactly equal to one when MSE optimal linear predictors are used for all bands. This can be seen by applying (16) to each of the M individual subbands and then multiplying the results together. To do this, we need an expression for the power spectral density of $y_k(m, n)$, the decimated subband signal. It is relatively easy to show that, for ideal filters, we have

$$S_{y_k}(\omega_1, \omega_2) = \frac{1}{M} S_{x_k}\left(\frac{\omega_1}{M}, \frac{\omega_2}{M}\right). \tag{18}$$

The factor $1/M$ is necessary to maintain equal variance between x_k and y_k. Inserting (18) in (16) and carrying out the previously indicated multiplication, we obtain

$$\left(\prod_{k=1}^{M}\sigma_{p,k}^2\right) = \frac{1}{M^M}\exp\left[\frac{M}{(2\pi)^2}\int_{-\pi}^{+\pi}\int_{-\pi}^{+\pi} \cdot \ln S_x(\omega_1, \omega_2)\,d\omega_1,\,d\omega_2\right] = \frac{1}{M^M}\,\sigma_p^{2M}.$$

Thus,

$$M\,\sigma_{gm}^2 = \sigma_p^2,$$

so that we have

$$G_{\text{SBC/DPCM}} = 1. \tag{19}$$

The coding gain of SBC relative to PCM is thus the same as that of DPCM relative to PCM

$$G_{\text{SBC/PCM}} = \gamma_y^{-2}. \tag{20}$$

Note that coding gains (17), (19), and (20) are for the case where the bit rate is not too small so that asymptotic formulas such as (12) can be used for the quantizer MSE. Also, the coding gains with respect to PCM assume Gaussian data (in cancelling out the quantizer efficiency factor $g(B)$ which is pdf dependent). Further, these for-

[1]Note: We can cancel the g's here because SBC and fullband DPCM both have the Laplacian pdf at their quantizers.

mulas assume optimal linear prediction. In a practical case, we would expect SBC to perform better than DPCM because: 1) restricted to a subband our low-order predictors will be more nearly optimal, and 2) although B is small, the B_k corresponding to bands with significant energy may be large, thus making (13) and (14) more realistic in the subband coding case.

Comments

1) If we took M large enough, we could efficiently use PCM to encode the subbands since their power spectral density would be nearly white. For more modest choices of M, several of the subbands will benefit from DPCM coding.

2) The gain for transform coding with the DFT or DCT is also given by the inverse of the spectral flatness measure in the high rate case [2]. Thus, the SBC coder can theoretically perform equally to the DCT coder at high bit rates for Gaussian images.

3) As M increases from one, the subband coder employing DPCM moves continuously from fullband DPCM towards full-image DFT encoding retaining asymptotic optimality in the sense of (20) at each step.

IV. Adaptive Subband Coding

In addition to bit allocation between the subbands, bits can also be allocated within the subbands, i.e., in the spatial domain. It is clear that within any image there are areas of large change such as edges, and areas of smaller change, as well as regions which are relatively flat. If we divide the image up into 16×16 blocks, corresponding to 4×4 blocks in each of the subband images, we can use the local variance to separate the blocks into busy, nonbusy, and quiet categories. We can then assign more bits to the busy blocks and fewer bits to the other blocks, tacitly accepting an overhead requirement of less than 0.01 bits/pixel. Choosing local variance thresholds to provide equal numbers of blocks in each category, we can see that the reconstruction error variance in subband k will be the average of the subband k error variances in each of the three categories. Denoting the error variance in category j as $\sigma_{k,j}^2$, we have

$$\sigma_r^2 = \tfrac{1}{3} \sum_{j=1}^{3} \sum_{k=1}^{M} \sigma_{k,j}^2.$$

Then substituting in (12) for each of the three busyness categories, we get

$$\sigma_r^2 = \tfrac{1}{3}\, g \sum_{k,j} \sigma_{p,k,j}^2\, 2^{-2B_{k,j}}.$$

This MSE is optimized using the same method as used on (13) to obtain the bit assignment

$$B_{k,j} = B + \frac{1}{2} \log_2 \left(\frac{\sigma_{p,k,j}^2}{\sigma_{gm,a}^2} \right), \qquad 1 \le k \le M, \quad 1 \le j \le 3, \tag{21}$$

where

$$\sigma_{gm,a}^2 \triangleq \left(\prod_{k,j} \sigma_{p,k,j}^2 \right)^{1/3M}.$$

A problem with (21) as well as (15) earlier is that the B_{kj} may not be integers, and they may (even worse) be negative! Actually, we only need $2^{2B_{kj}}$ to be an integer, and this can often be handled by approximation. The negative values are more of a problem, but it is a standard one in the area of transform coding. Quickly converging iterative algorithms have been developed to deal with negative B_{kj} by truncating the negative values and then evenly subtracting off the excess average bit rate thus created [2]. Generally no more than two iterations of this process are necessary.

Within each category and each subband, we solve the least-squares prediction problem to find the appropriate AR model. The error variance within each category $\sigma_{p,k,j}^2$ then permits the selection of a Laplacian optimized quantizer [12] for the assigned bit rate and, hence, the number of levels.

V. Coding Simulation Results

A coding simulation for SBC was carried out at the Image Processing Laboratory at R.P.I. We considered two monochrome images of size 256×256 with 8-bit gray scales. The encoding was done in the density or log-intensity domain. The images are referred to as "Lady" and "Building" and are seen in Figs. 12(a) and 15, respectively. Both images were encoded at $B = 2.0$ and 1.0 bits per pixel (bpp). The "Lady" image was also encoded at $B = 0.67$ bpp, while the "Building" image was encoded at $B = 0.57$ bpp. The "Lady" image was encoded both adaptively and nonadaptively, however, the "Building" image was only coded adaptively. Convolution was performed circularly with $N = 256$.

For $B = 1.0$ and the "Lady" image, adaptive bit allocation gave the results shown in Table I. These numbers reflect the iterative adjustment to remove negative values. Using normalized Laplacian quantizers with numbers of levels 2, 3, 4, 5, 8, 16, and 32, these values were then modified to the bit assignments shown in Table II. Here the numbers in parentheses indicate 3 and 5 level "midtread" quantizers (i.e., representation levels at zero) which were coded by short length Huffman codes as seen in Table III with the resulting average wordlength as seen in Table II.

Fig. 12 shows the coding results for the "Lady" image using a nonadaptive SBC approach with 16 subbands and 1×1-order quarter-plane predictors. Fig. 12(b) was coded at 0.67 bpp, Fig. 12(c) was coded at 1.0 bpp, and Fig. 12(d) was coded at 2.0 bpp. Fig. 13 shows the corresponding result for adaptive SBC. Fig. 14 shows closeups of the center portion of the adaptive SBC images of Fig. 13. Viewing the original images we discern a slight impairment at 2.0 bpp and a quite noticeable impairment of 0.67 bpp, with the 1.0 bpp result being in between. Fig.

TABLE I
ADAPTIVE BIT ALLOCATIONS FOR THE SUBBANDS OF "LADY" USING (21) AND B = 1 bpp (ALL VALUES HAVE BEEN ADJUSTED TO ELIMINATE NEGATIVE BIT ASSIGNMENTS)[a]

	Sub. 11-11	Sub. 11-12	Sub. 11-21	Sub. 11-22
Quiet	3.68	1.61	0.89	0.76
Nonbusy	4.54	2.65	1.79	1.79
Busy	4.71	3.57	2.58	2.61
	Sub. 12-11	Sub. 12-12	Sub. 12-21	Sub. 12-22
Quiet	0.00	0.86	0.00	0.00
Nonbusy	1.05	1.91	0.43	0.61
Busy	1.75	2.21	1.31	1.96
	Sub. 21-11	Sub. 21-12	Sub. 21-21	Sub. 21-22
Quiet	0.00	0.00	0.00	0.00
Nonbusy	0.00	0.00	0.00	0.10
Busy	0.31	0.21	1.05	1.08
	Sub. 22-11	Sub. 22-12	Sub. 22-21	Sub. 22-22
Quiet	0.00	0.00	0.00	0.00
Nonbusy	0.00	0.00	0.00	0.22
Busy	0.30	0.39	0.48	1.17

[a]Sub. *ij-kl* refers to the subband resulting from filtering the image by H_{ij} and then by H_{kl}.

TABLE II
ACTUAL BIT ASSIGNMENTS FOR B = 1 bpp ON "LADY." ASSIGNMENTS WERE CALCULATED BY ROUNDING THE VALUES IN TABLE I TO THE CLOSEST POSSIBLE BIT ASSIGNMENTS. ALLOCATIONS OF 1 AND 2 BITS WERE REPLACED BY 3-LEVEL AND 5-LEVEL QUANTIZERS FOR SUBBANDS OTHER THAN 11-*xx*

Sub. 11-11	Sub. 11-12	Sub. 11-21	Sub. 11-22
4	2	1	1
5	3	2	2
5	4	3	3
Sub. 12-11	Sub. 12-12	Sub. 12-21	Sub. 12-22
0	1.36 (3)	0	0
1.17 (3)	1.86 (5)	0	0
1.28 (3)	1.85 (5)	0	1.76 (5)
Sub. 21-11	Sub. 21-12	Sub. 21-21	Sub. 21-22
0	0	0	0
0	0	0	0
0	0	1.10 (3)	1.20 (3)
Sub. 22-11	Sub. 22-12	Sub. 22-21	Sub. 22-22
0	0	0	0
0	0	0	0
0	0	0	1.17 (3)

Total = 47.75/48 = 1.00 bit

TABLE III
VARIABLE LENGTH CODE FOR 3 AND 5 RECONSTRUCTION LEVELS

Reconstruction Level	Code	Code Length
−1.422	10	2
0.0	0	1
1.422	11	2
−2.254	100	3
−0.840	101	3
0.0	0	1
0.840	110	3
2.254	111	3

Reconstructioan levels are normalized.

Fig. 12. Nonadaptive SBC results for "Lady" image: (a) original, (b) 0.67 bpp, (c) 1.0 bpp, (d) 2.0 bpp.

15 shows the corresponding adaptive SBC coding results for the "Building" image. Here the lowest rate is 0.57 bpp. Visual impairments are not evident in the 1.0 and 2.0 bpp "Building" images, however, the 0.57 bpp image seems to have a slightly "muddy" quality. Fig. 16 shows an SNR comparison for adaptive and nonadaptive SBC to three other existing techniques extending the comparison presented in [10]. The other curves plotted are taken from [10] and are the SNR performance of the adaptive DCT of Chen and Smith [9] and two vector quantization approaches. We see that the adaptive SBC gives the best SNR performance at all bit rates in the range 0.67–2.0. Also, the nonadaptive SBC is very competitive

Fig. 13. Adaptive SBC results for "Lady" image: (a) original, (b) 1.0 bpp, (c) 0.67 bpp, (d) 2.0 bpp.

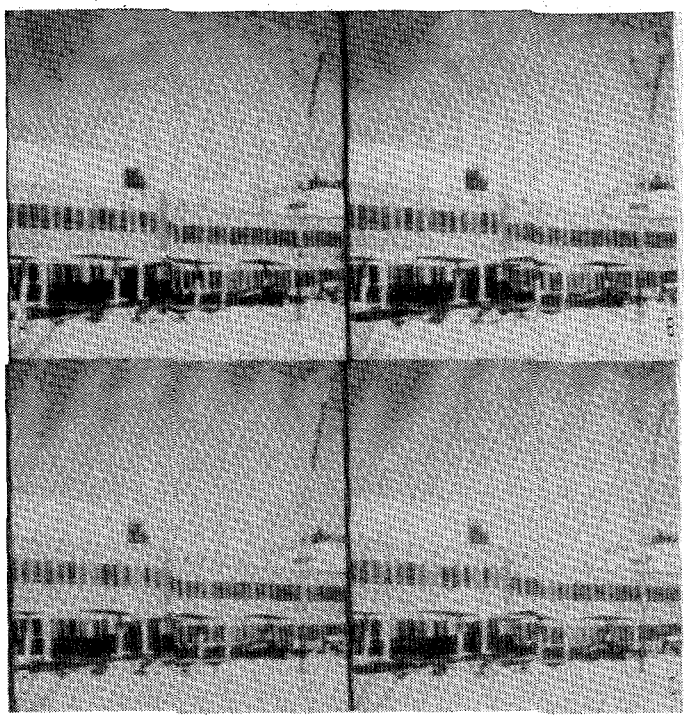


Fig. 15. Adaptive SBC results for "Building" image: (a) original, (b) 0.57 bpp, (c) 1.0 bpp, (d) 2.0 bpp.

Fig. 14. Closeups of adaptive SBC results of Fig. 13: (a) original, (b) 0.67 bpp, (c) 1.0 bpp, (d) 2.0 bpp.

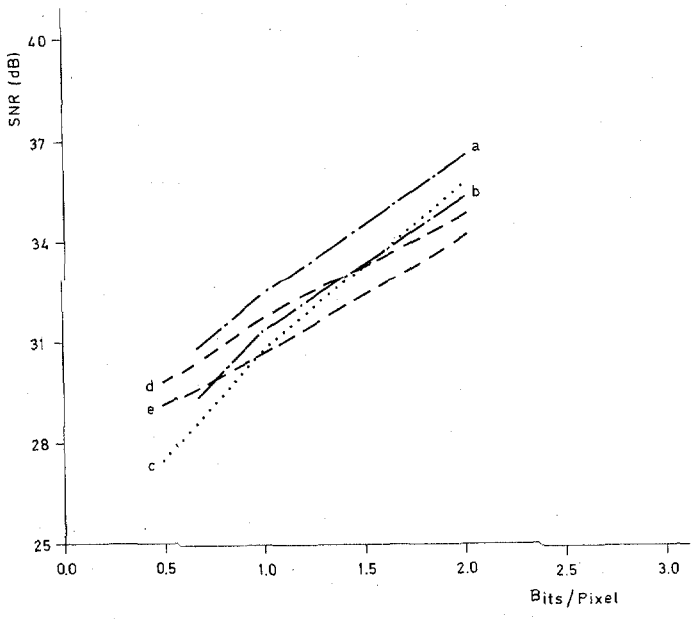


Fig. 16. SNR versus bit rate for "Lady" image using: (a) *adaptive SBC*, (b) nonadaptive SBC, (c) adaptive DCT, (d) differential VQ, and (e) VQ. (Plots of (c), (d), and (e) are taken from [10].)

with the other techniques over this range. Table IV lists our SNR results for SBC image coding.

VI. Conclusions

In this paper we have analyzed a new method of image coding called subband coding. This method has several advantages. First, it has the best SNR performances of the methods tested. Second, subband coding has good subjective error properties. Third, complexity comparable to transform coders is achievable using pseudo-QMF's. Subband image coding facilitates the progressive transmission approach, whereby lower resolution data are sent first and the detail appears gradually when the image does not change. The SBC technique is compatible with respect to high- and low-definition systems, in that the low-definition receivers can just ignore the higher resolution subbands. The new coding technique seems to have the performance of transform coding at low rates without the attendant blocky structure. Finally, one can see that subband coding unifies transform coding and DPCM by pro-

TABLE IV
SNR VERSUS BIT RATES FOR "LADY" AND "BUILDING"

Bits	"Lady" Nonadaptive	"Lady" Adaptive	"Building" Adaptive
0.67	29.4 dB	30.9 dB	30.6 dB[a]
1.0	31.4 dB	32.5 dB	33.8 dB
2.0	35.4 dB	36.6 dB	38.0 dB

[a]0.57 bits.

viding a continuum between these two extremes indexed by the parameter M, the number of subbands.

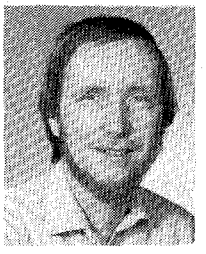
Suggestions for further study include the following. As in 1-D subband coding of speech, it should be possible to shape the error spectrum to better match the human visual sensitivity. The separable QMF structure is efficient but it is also restrictive in terms of the directional orientation of the subbands. For example, one may want to split the 22 subband into 45° and 135° subbands in order to use more directional predictors in the DPCM coders. This cannot be done with separable filters. However, a nonseparable pseudo-QMF structure might exist to solve this problem.

ACKNOWLEDGMENT

The authors would like to thank an anonymous reviewer for pointing out a mistake relating to cancellation of the aliasing terms in (3) in an earlier version of this paper.

REFERENCES

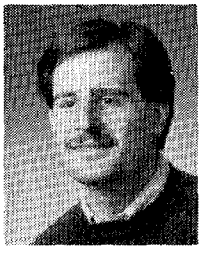
[1] R. E. Crochiere, S. A. Webber, and J. L. Flanagan, "Digital coding of speech in subbands," *Bell. Syst. Tech. J.*, vol. 55, pp. 1069–1085, Oct. 1976.
[2] N. S. Jayant and P. Noll, *Digital Coding of Waveforms.* Englewood Cliffs, NJ: Prentice-Hall, 1984.
[3] E. R. Kretzmer, "Reduced-alphabet representation of television signals," *IRE Convention Rec.*, vol. 4, pp. 140–147, 1956.
[4] W. F. Schreiber, C. F. Knapp, and N. D. Kay, "Synthetic highs, An experimental TV bandwidth reduction systems," *J. SMPTE*, vol. 68, pp. 525–537, Aug. 1959.
[5] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM-31, pp. 532–540, Apr. 1983.
[6] S. D. O'Neil, "Sub-band coding of images with adaptive bit allocation," M.S. thesis, ECSE Dep., R.P.I., Troy, NY, Apr. 1985.
[7] M. Vetterli, "Multi-dimensional sub-band coding: Some theory and algorithms," *Signal Processing*, vol. 6, pp. 97–112, Apr. 1984.
[8] D. Esteban and C. Galand, "Application of quadrature mirror filters to split band voice coding schemes," in *Proc. ICASSP*, May 1977, pp. 191–195.
[9] W. H. Chen and C. H. Smith, "Adaptive coding of monochrome and color images," *IEEE Trans. Commun.*, vol. COM-25, pp. 1285–1292, Nov. 1977.
[10] R. L. Baker and R. M. Gray, "Image compression using non-adaptive spatial vector quantization," in *Proc. 16th Asilomar Conf.*, Nov. 1982.
[11] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," in *Proc. ICASSP*, Apr. 1980, pp. 291–294.
[12] K. Nitadori, "Statistical analysis of DPCM," *Electron. Commun. Japan*, vol. 48, pp. 17–26, Feb. 1965.



John W. Woods (S'67–M'70–SM'83) was born in Washington, DC, on December 5, 1943. He received the B.S., M.S., E.E., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1965, 1967, and 1970, respectively.

From 1970 to 1973 he was at the VELA Seismological Center, Alexandria, VA, working on array processing of digital seismic data. From 1973 to 1976 he was at the Lawrence Livermore Laboratory, University of California, Berkeley, working on two-dimensional digital signal processing. Since 1976 he has been with the Department of ECSE at Rensselaer Polytechnic Institute, Troy, NY, where he is currently a Professor. He has taught courses in digital signal processing, probability and stochastic processes, information theory, and communication systems. His research interests include estimation/restoration, detection, recursive digital filtering, and data compression of images and other multidimensional data. He has authored or coauthored over 40 papers in these fields. During the academic year 1985–1986, he was Visiting Professor in the Information Theory Group at Delft University of Technology, Delft, The Netherlands.

Dr. Woods was co-recipient of the 1976 Senior Award of the ASSP Society. He is a former member of the Digital Signal Processing Committee. He was Chairman of the Schenectady joint ASSP/Communications Society Chapter in 1977–1978. He is a former Associate Editor for Signal Processing of the IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. He was Co-chairman of the Third ASSP Workshop on Multidimensional Signal Processing held at Lake Tahoe, CA, in October 1983. He is a former Chairman of the ASSP Technical Committee on Multidimensional Signal Processing. He is a member of Sigma Xi, Tau Beta Pi, Eta Kappa Nu, and the AAAS. He is currently a member of the ASSP AdCom.



Sean D. O'Neil was born in Chicago, IL, on August 30, 1961. He received the B.S. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1983, and the M.S. degree from the Rensselaer Polytechnic Institute, Troy, NY, in 1985.

Since 1985 he has been working at the MITRE Corporation, Bedford, MA, on number theoretic signal processing and its implementation in VLSI. His interests include image compression, statistical pattern recognition, machine vision, and their applications.