

# Orientation Adaptive Subband Coding of Images

David Taubman and Avidesh Zakhor, *Member, IEEE*

**Abstract**—In the subband coding of images, directionality of image features has thus far been exploited very little. The proposed subband coding scheme utilizes orientation of local image features to avoid the highly objectionable Gibbs-like phenomena observed at reconstructed image edges with conventional subband schemes at low bit rates. At comparable bit rates, the subjective image quality obtained by our orientation adaptive scheme is considerably enhanced over a conventional separable subband coding scheme, as well as other separable approaches such as the JPEG compression standard.

## I. INTRODUCTION

IN recent times, interest has been developed in adaptation of the parameters of filter bank systems used for subband coding. Most work in this area thus far has concentrated on the adaptation of 1-D filter coefficients. Delsarte *et al.* [8] have developed a “ring algorithm” to locally optimize the coefficients of a 1-D, two channel, paraunitary subband system for minimum mean squared error in a coding scenario. In [17] and [22], Taubman and Zakhor consider globally optimal solutions to the same problem and consider the adaptation of certain nonseparable 2-D filter systems, given a known decimation structure. Gurski *et al.* [10] solve a related problem in the 1-D case. They consider pyramidal decompositions rather than subband decompositions and adapt the analysis and synthesis filters in an iterative way to minimize the signal energy in the difference channel. Nayebe *et al.* [15], on the other hand, take a different approach in which they maintain a library of preoptimized compatible subband filters from which an optimal selection is made. Their work tackles the issue of preserving perfect reconstruction when the filter impulse responses vary with location as they must in an adaptive environment. This work, however, is also restricted to one dimension.

In a multidimensional setting, directionality arises as a new and important attribute that is absent from 1-D systems. Attempts to exploit this property in image compression may be classified into two broad categories. In the first category are those strategies that analyze an image along a predetermined set of directions. In the second are strategies that adapt the directional analysis itself to orientational features of the image being analyzed. By far, the majority of approaches currently proposed to exploit image directionality come under the first category.

Manuscript received December 16, 1992; revised August 4, 1993. The associate editor coordinating the review of this paper and approving it for publication was Prof. William A. Perlman.

The authors are with the Department of Electrical Engineering and Computer Science, University of California at Berkeley, Berkeley, CA 94720, USA.

IEEE Log Number 9400314.

Image compression strategies belonging to the first category are more or less based on studies of the mammalian visual cortex. These studies indicate that receptive fields of cortical cells have definite orientational tuning with average orientational bandwidths of approximately  $35^\circ$ . Daugman [6], [7] applies such psychovisual evidence to a model of the human visual cortex based on 2-D Gabor functions. He demonstrates an image compression scheme in which a neural network is used to represent images in terms of the nonorthogonal set of 2-D Gabor basis functions, each of which has an orientational bandwidth of  $30^\circ$ . A related, so-called “cortical transform” is proposed by Watson [24]. His technique is computationally simpler but leads to many more transform coefficients than original image samples, unlike Daugman’s scheme, which preserves sample count. Rao and Pearlman [16] propose an image compression scheme based on subband techniques in which Daugman’s ideal 2-D Gabor filters are approximated with so-called “dome” filters. Unfortunately, their scheme also leads to some expansion in the number of samples that must be coded.

A number of nonseparable maximally decimated perfect reconstruction subband systems with psychovisually meaningful orientational sensitivity may also be classified under the first category. Such systems have been proposed by Mahesh and Pearlman [13], Simoncelli and Adelson [20], and Bamberger and Smith [2], [3]. The schemes of Mahesh, Pearlman, Simoncelli and Adelson are based on hexagonally sampled images and lead to subbands with more cortically relevant orientational sensitivities than conventional separable subband systems. Bamberger and Smith, on the other hand, demonstrate perfect reconstruction filter banks, based on a conventional rectangular image sampling, capable of resolving images into many different directional components. Chang and Zakhor [4] have also proposed a nonseparable 3-D subband scheme to take advantage of the spatio-temporal directionality of video sequences by decomposing them into different velocity components.

We also make mention of the image coding scheme proposed by Ikonomopoulos and Kunt [11], which uses a fixed set of directional filters to identify oriented edge segments, for which an efficient coding scheme is developed. Their scheme allows an image to be approximately recovered from the coded edge information for applications requiring very high compression ratios. Moreover, Li and He [12] propose an image compression algorithm that combines subband decomposition with Ikonomopoulos’ and Kunt’s scheme for identifying and coding oriented edge segments.

All of the above-mentioned approaches fall into the first category of attempts to exploit directionality in image com-

pression. They each involve psychovisually relevant yet fixed directional image decompositions. Such schemes may be considered adaptive in the sense that adaptive quantization and coding schemes could be applied to the directional components. To our knowledge, however, few authors propose image compression algorithms in the second category, in which the directional decomposition itself is adaptive. In addition to the work presented in this paper, that of Shapiro [18] may be classified under this second category. He proposes a McClellan transform-based approach to the directional adaptation of nonseparable multidimensional filter banks with the well-known two-channel quincunx subsampling structure.

In this paper, we propose an adaptive subband scheme that directly exploits the directional attributes of images. Our work is based on the observation that natural images commonly contain approximately linear edges, on a local level at least, whose orientation is neither vertical nor horizontal, the directions in which a separable subband approach performs best. Such features result in considerable activity in the high-frequency subbands of a conventional separable subband scheme. In addition, at low bit rates, the quantization effects can be observed strongly at such edges in the form of Gibbs phenomena [19], giving them a ragged, ringing appearance, which becomes highly objectionable as the bit rate is reduced.

If, however, the image is first resampled along axes aligned to its dominant linear features, then in the resampled domain, these features become vertical and/or horizontal; of course, in so doing, the image's region of support ceases to be rectangular. A conventional separable subband system may then be applied in the resampled domain, resulting in reduced activity in the high-frequency subbands. Moreover, when the subbands are quantized, artifacts induced in the reconstructed image tend to lie parallel to the dominant linear image features, these being oriented vertically and horizontally in the resampled domain. Subjectively, these artifacts tend to be much less disturbing than those generated by the conventional subband scheme. These phenomena are illustrated in the synthetic examples presented in Section III and the compression results presented in Section VI.

Fig. 1 shows the analysis system of our proposed directional subband scheme. The system is the cascade of an image resampling network and a conventional four-channel separable analysis system with  $F$  a low-pass and  $G$  a high-pass filter. The synthesis system is a conventional separable synthesis system followed by an inverse resampling network. Only one level of the anticipated pyramid structure, which is obtained by branching along the LL channels, is shown in the figure. In our experimental work, for example, we find three levels to give satisfactory coding performance. This particular case results in nine high-frequency subbands and one low-pass subband in all.

In overview, our proposed algorithm is as follows. We first partition an image into smaller rectangular regions, which will be considered independently. This is necessary to minimize the number of distinct orientations presented to the adaptive subband system. For each partition, we then choose a pair of axes along which to resample. These axes are chosen to correspond to the dominant linear features in the partition. The partition is then resampled along these axes, following which,

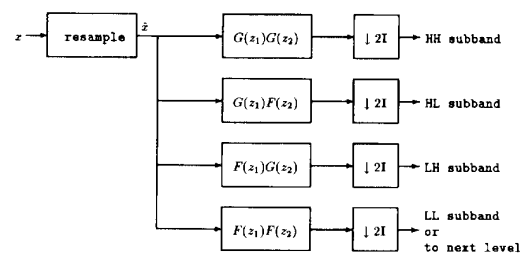


Fig. 1. Directional analysis system schematic.

a conventional separable subband decomposition is applied in the resampled domain. The subband samples are quantized and run-length and entropy coded to be transmitted along with the orientation information.

At the receiver, each partition is reconstructed in the resampled domain, and then, using the orientation information, the resampling process is inverted. Finally, to remove artifacts resulting from independent processing of the image partitions, a local smoothing algorithm is applied at the partition boundaries. This algorithm also exploits directional attributes of the image.

In the remainder of the paper, we will elaborate on the various aspects of this algorithm. In Section II, we discuss the resampling process mentioned above, by which we accomplish orientational adaptivity. In Section III, we present synthetic examples that illustrate the performance of our resampling scheme in conjunction with a separable subband coder. In Section IV, we present our algorithm for the choice of optimum resampling axes for each image partition, which relies on image feature extraction techniques. The proposed algorithm for removing partitioning artifacts in reconstructed images appears in Section V. Our choice of subband filters, the partitioning strategy, and details of our quantization and coding approach are all provided in Section VI. In Section VI, we also present a comparison of reconstructed images using conventional and orientation adaptive subband coders. Then, in Section VII, we discuss the complexity of our proposed scheme.

## II. RESAMPLING

This section is devoted to discussion of the resampling block in Fig. 1, by which we achieve orientational tuning of our otherwise separable subband coding system. We begin, in Section II-A, by considering the reorientation of spatially continuous signals as a vehicle for analyzing the resampling of our discrete space signals. The receiver for our subband system must be able to invert the resampling process in order to reconstruct the image; accordingly, Section II-B is concerned with invertible resampling of discrete space signals in a manner consistent with the reorientation of spatially continuous signals outlined in Section II-A. We proceed to discuss a subset of such invertible resampling operators, which may be implemented by index permutation and interpolation operations. We discuss implementation of the interpolation in Section II-C. Based on these considerations, we propose a two-parameter resampling

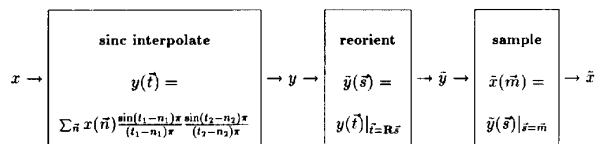


Fig. 2. Ideal resampling of discrete space signals.

system in Section II-D, the parameters of which are related to the two resampling axes. Finally, in Section II-E, we present a pictorial explanation of the operation of the resampling block in Fig. 1.

#### A. Reorientation of Spatially Continuous Signals

Let  $y(\vec{t})$  denote a 2-D signal of spatially continuous variable  $\vec{t}$ , bandlimited to the open region  $(-\pi, \pi)^2$ , and let  $\mathbf{R}$  denote a nonsingular matrix that we will later refer to as the resampling matrix. We define our reorientation operation by

$$\tilde{y}(\vec{s}) = y(\vec{t}) \quad \text{with} \quad \vec{t} = \mathbf{R}\vec{s} \quad (1)$$

where  $\tilde{y}(\vec{s})$  is the reoriented signal of spatially continuous variable  $\vec{s}$ . In the Fourier domain, (1) becomes

$$\tilde{Y}(\vec{\psi}) = \det(\mathbf{R}^{-1})Y(\vec{\omega}) \quad \text{where} \quad \vec{\omega} = \mathbf{R}^{-t}\vec{\psi}. \quad (2)$$

Noting that  $Y(\vec{\omega})$  has region of support  $\vec{\omega} \in (-\pi, \pi)^2$ , (2) implies that  $\tilde{Y}(\vec{\psi})$  has region of support  $\mathcal{W}_{\tilde{y}}$ , the region bounded by the parallelogram with vertices at  $\vec{\psi} = \mathbf{R}^t \cdot [\pm\pi, \pm\pi]^t$ .

For our orientation adaptive subband scheme, we are interested in matrices  $\mathbf{R}$ , which reorient an image. A natural choice would be to rotate the image, in which case,  $\mathbf{R}$  would be an orthogonal matrix. Despite the intuitive plausibility of such a choice, we shall see in the next section that this does not lead to an invertible resampling operator. Instead of rotation, we will use matrices  $\mathbf{R}$ , which skew the image to achieve our reorientation.

#### B. Invertible Resampling of Discrete Space Signals

The objective of the resampling block, in our orientation adaptive subband scheme of Fig. 1, is to reorient the signal so that in the resampled domain, dominant linear image features appear oriented in the cardinal directions—i.e., vertically and horizontally. We assume that the 2-D sequence  $x$  in Fig. 1 is only a sampling of an underlying spatially continuous signal  $y$  bandlimited to  $(-\pi, \pi)^2$ . Ideally, the resampling operation desired would recover  $y$  from  $x$  by sinc interpolation; reorient the spatially continuous image  $y$  according to (1) with some resampling matrix  $\mathbf{R}$ , and then, sample the reoriented image  $\tilde{y}$  to obtain the new sequence  $\tilde{x}$ . This sequence of operations is depicted in Fig. 2. The remainder of this section is devoted to the two issues of invertibility and implementation feasibility, which influence our ability to approximate this ideal resampling.

Intuitively, the requirement that our resampling operation be invertible means that the sequence  $x$  should always be recoverable from the resampled sequence  $\tilde{x}$ . Mathematically,

we consider discrete space resampling as a linear operator  $\mathcal{R} : l^1 \rightarrow l^1$ , where  $l^1$  is the set of absolutely summable sequences—2-D sequences in our case. We require  $\mathcal{R}$  to be one-to-one and conservative of sample density.<sup>1</sup> Sample density conservation is clearly important in coding applications. The reason for imposing the requirement that  $\mathcal{R}$  be one-to-one is that otherwise, even in the absence of quantization error, the receiver would be unable to recover  $x$  from  $\tilde{x}$ , and therefore, our subband system could not be considered perfect reconstruction.

As noted in Section II-A,  $\tilde{Y}(\vec{\psi})$  has region of support  $\mathcal{W}_{\tilde{y}}$ , which is the region bounded by the parallelogram with vertices at  $\vec{\psi} = \mathbf{R}^t \cdot [\pm\pi, \pm\pi]^t$ . If the cosets  $\mathcal{W}_{\tilde{y}}^{\vec{k}}$  given by  $\mathcal{W}_{\tilde{y}}^{\vec{k}} = \mathcal{W}_{\tilde{y}} + 2\pi\vec{k}$ , where  $\vec{k} \in \mathcal{Z}^2$ , are not disjoint, then aliasing will occur in sampling  $\tilde{y}$  to obtain  $\tilde{x}$ . In this case, the resampling operator  $\mathcal{R}$  depicted in Fig. 2 cannot be one to one. On the other hand, if the cosets  $\mathcal{W}_{\tilde{y}}^{\vec{k}}$  do not cover the entire spectrum, then each coset must have area less than  $4\pi^2$ ; but, from (2), the area of  $\mathcal{W}_{\tilde{y}}$  is  $4\pi^2 \cdot |\det(\mathbf{R})|$ . Thus, we would have  $|\det(\mathbf{R})| < 1$ , and the sampling density would be increased, violating our requirement of sample density conservation.

In summary, a requirement for invertible resampling is that  $\mathbf{R}$  be chosen in such a way that the cosets  $\mathcal{W}_{\tilde{y}}^{\vec{k}}$  tessellate the frequency spectrum. We note here that although orthogonal matrices do have  $\det(\mathbf{R}) = 1$ , they do not in general lead to cosets  $\mathcal{W}_{\tilde{y}}^{\vec{k}}$ , which tessellate the frequency spectrum; therefore, image rotation is excluded as an invertible resampling.<sup>2</sup> A simple and, as we shall see, numerically efficient means of satisfying the tessellation constraint is to build  $\mathbf{R}$  as a product of matrices of the form

$$\mathbf{R}_\alpha = \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} \quad \text{or} \quad \mathbf{R}_\beta = \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix}. \quad (3)$$

These correspond to a skewing of the image. Observing that a resampling of the form  $\mathbf{R}_\beta$  may be implemented by transposing the image and then applying a resampling of the form  $\mathbf{R}_\alpha$ , we will only consider the latter.

Applying  $\mathbf{R}_\alpha$  to  $y$ , we obtain

$$\tilde{y}(s_1, s_2) = y(s_1, s_2 + \alpha s_1).$$

Combining this with the sampling of  $\tilde{y}$  in Fig. 2 yields

$$\tilde{x}(m_1, m_2) = y(t_1, t_2)|_{t_1=m_1, t_2=m_2+\alpha m_1} \quad (4)$$

which is satisfied by the following two relations that correspond to building blocks in our resampling algorithm: an interpolative block satisfying

$$\begin{aligned} \tilde{x}(l_1, l_2) &= y(t_1, t_2)|_{t_1=l_1, t_2=l_2+\sigma(l_1)} \\ \text{where } \sigma(l_1) &= \alpha l_1 - \langle \alpha l_1 \rangle \end{aligned} \quad (5)$$

<sup>1</sup>We use the intuitive expression ‘‘sample density conservation’’ to indicate that all 2-D regions of support  $A$  are mapped by  $\mathcal{R}$  to sequences with bounded region of support  $B$  such that  $\frac{\#(B)}{\#(A)} \rightarrow 1$  as  $\#(A) \rightarrow \infty$  with  $\#(A)$  and  $\#(B)$  denoting the number of points in these regions of support. Given the sequence of operations depicted in Fig. 2, this condition is clearly satisfied if and only if  $|\det(\mathbf{R})| = 1$ .

<sup>2</sup>This fact may come as a surprise to the reader accustomed to working with matrices with integral coefficients, for which  $|\det(\mathbf{R})| = 1$  is sufficient for invertibility.

in which  $\langle \cdot \rangle$  denotes rounding to the nearest integer and an index permutation block satisfying

$$\tilde{x}(m_1, m_2) = \tilde{x}(l_1, l_2)|_{l_1=m_1, l_2=m_2+\langle \alpha m_1 \rangle}. \quad (6)$$

Referring to Fig. 2,  $y$  is to be obtained from  $x$  by sinc interpolation; therefore, (5) becomes

$$\begin{aligned} \tilde{x}(l_1, l_2) &= \sum_{n_2} x(l_1, n_2) \frac{\sin(l_2 + \sigma(l_1) - n_2)\pi}{(l_2 + \sigma(l_1) - n_2)\pi} \\ &= \sum_{n_2} x(l_1, n_2) f_{\sigma(l_1)}(l_2 - n_2) \end{aligned} \quad (7)$$

where

$$f_{\sigma(j)}(i) = \frac{\sin(i + \sigma(j))\pi}{(i + \sigma(j))\pi}. \quad (8)$$

Observe, then, that a resampling corresponding to  $\mathbf{R}_\alpha$ , or equivalently  $\mathbf{R}_\beta$ , involves only the 1-D filtering and index permutation described by (7) and (6), respectively. In this setting, invertibility of our resampling operator depends only on the invertibility of this filtering. In the next section, we discuss means to approximate the interpolative filtering of (7) with an invertible operator. We show that although the interpolation must necessarily be approximated, invertibility may be achieved exactly.

### C. Interpolation

In this section, we discuss an approximation of the interpolator idealized by (7) and (8), that is, our approximate interpolator maps sequences  $x$  to sequences  $\tilde{x}'$ , where

$$\tilde{x}'(l_1, l_2) \approx \tilde{x}(l_1, l_2).$$

We first observe that the infinite extent impulse response  $f_{\sigma(j)}$  of (8) must be windowed to produce realizable finite extent filters. In particular, we arbitrarily choose to employ a Hanning window to  $f_{\sigma(j)}$  to obtain the approximate interpolation filter

$$\begin{aligned} \hat{f}_{\sigma(j)}^N(i) &= \begin{cases} \frac{1}{2} f_{\sigma(j)}(i) \left[ 1 + \cos\left(\frac{i+\sigma(j)}{\tau}\pi\right) \right] & , |i + \sigma(j)| < \tau \\ 0 & , \text{otherwise} \end{cases} \\ &= \begin{cases} \frac{1}{2} \frac{\sin(i+\sigma(j))\pi}{(i+\sigma(j))\pi} \left[ 1 + \cos\left(\frac{i+\sigma(j)}{\tau}\pi\right) \right] & , |i + \sigma(j)| < \tau \\ 0 & , \text{otherwise} \end{cases} \end{aligned}$$

where

$$\tau = N + 1 - \sigma(j).$$

Noting that  $|\sigma(j)| \leq \frac{1}{2}$  by (5), these filters  $\hat{f}_{\sigma(j)}^N(i)$  have support  $-N \leq i \leq N$ .

In practice, we must work with finite extent sequences. Zero padding would introduce highly undesirable transients at the boundaries. We choose, instead, to handle boundaries by reducing the window size  $N$  of our interpolation filters  $\hat{f}_{\sigma(j)}^N$ , as image boundaries are approached.<sup>3</sup> To be precise, we

<sup>3</sup>An alternative approach would be to keep the window size  $N$  fixed and to symmetrically extend [21] the image boundaries. We find that such an approach adversely affects the rate of convergence of the iterative inversion algorithm described shortly.

define the 1-D interpolation operators  $\mathcal{S}_{\sigma(j)} : l^1 \rightarrow l^1$  by

$$\mathcal{S}_{\sigma(j)}(\varphi) = \tilde{\varphi}$$

where

$$\tilde{\varphi}(n) = \sum_{k=-N(n)}^{N(n)} \hat{f}_{\sigma(j)}^{N(n)}(k) \varphi(n-k)$$

and

$$N(n) = \min(|n - K_1|, |n - K_2|, N_{\max}). \quad (9)$$

Here,  $\varphi$ , and hence  $\tilde{\varphi}$ , have support  $[K_1, K_2]$ , and  $N_{\max}$  is chosen to limit the maximum interpolation filter size.  $\mathcal{S}_{\sigma(j)}$  corresponds to linear shift invariant (LSI) filtering everywhere except in the vicinity of  $K_1$  and  $K_2$ . It only approximately shifts discrete space sequences by  $\sigma(j)$  because the ideal shifting filter of (8) has been windowed.

We could use the operators  $\mathcal{S}_{\sigma(j)}$  to approximate the right-hand side of (7) with

$$\mathcal{S}_{\sigma(l_1)}(x(l_1, \cdot)). \quad (10)$$

However, the inversion of  $\mathcal{S}_{\sigma(j)}$  becomes an increasingly ill-conditioned problem as  $\sigma(j)$  approaches  $\pm\frac{1}{2}$ . This is a serious difficulty as  $\sigma(j)$  can take on any value in the interval  $[-\frac{1}{2}, \frac{1}{2}]$ .

That inversion of  $\mathcal{S}_{\sigma(j)}$  becomes ill-conditioned as  $\sigma(j)$  approaches  $\pm\frac{1}{2}$  may be understood by considering infinite extent sequences— i.e.,  $[K_1, K_2] = [-\infty, \infty]$ . In this case  $\mathcal{S}_{\sigma(j)}$  corresponds to LSI filtering by  $\hat{f}_{\sigma(j)}^{N_{\max}}$ . When  $|\sigma(j)| < \frac{1}{2}$ , half of the  $2N_{\max}$  zeroes of this FIR filter lie inside the unit circle, and the other half lie outside the unit circle; therefore, inversion may be accomplished by cascading a causally stable recursive filter and an anticausally stable recursive filter, where each has  $N_{\max}$  poles. The combined impulse response of this cascade is a two-sided IIR filter with absolutely summable impulse response, and hence,  $\mathcal{S}_{\sigma(j)}$  is invertible in the sense defined in Section II-B. At  $|\sigma(j)| = \pm\frac{1}{2}$ , however, one of the zeroes of  $\hat{F}_{\sigma(j)}^{N_{\max}}(z)$ , which is the Z-transform of  $\hat{f}_{\sigma(j)}^{N_{\max}}$ , crosses the unit circle at  $z = -1$ . Therefore,  $\mathcal{S}_{\pm\frac{1}{2}}$  is not invertible. It follows that inversion of  $\mathcal{S}_{\sigma(j)}$  becomes ill-conditioned as  $\sigma(j)$  approaches  $\pm\frac{1}{2}$ .

To resolve this difficulty, we choose to approximate (7) as

$$\tilde{x}(l_1, \cdot) \approx \tilde{x}'(l_1, \cdot) = \mathcal{S}_{-\frac{1}{2}\sigma(l_1)}^{-1} \circ \mathcal{S}_{\frac{1}{2}\sigma(l_1)} x(l_1, \cdot) \quad (11)$$

where  $\mathcal{S}_{-\frac{1}{2}\sigma(j)}^{-1}$  is the inverse of the operator  $\mathcal{S}_{-\frac{1}{2}\sigma(j)}$ .  $\mathcal{S}_{-\frac{1}{2}\sigma(j)}$  does not suffer from the ill-conditioning discussed above because  $\frac{1}{2}\sigma(j) \in [-\frac{1}{4}, \frac{1}{4}]$ . Equation (11) may be understood by recognizing that  $\mathcal{S}_{\frac{1}{2}\sigma(j)}$  approximates an ideal shift of its operand sequence by  $\frac{1}{2}\sigma(j)$  of a pixel separation. Thus,  $\mathcal{S}_{-\frac{1}{2}\sigma(j)}^{-1}$  inverts an approximate shift by  $-\frac{1}{2}\sigma(j)$  of a pixel separation. Hence, the composed operator in (11) performs an approximate shift by  $\frac{1}{2}\sigma(j) - (-\frac{1}{2}\sigma(j)) = \sigma(j)$  of a pixel separation.

In order to satisfy the requirement that our resampling be exactly invertible, it is necessary to accurately compute  $\mathcal{S}_{\frac{1}{2}\sigma(j)}^{-1} \tilde{\varphi}$  for any sequence  $\tilde{\varphi}$ . Our algorithm for implementing the operator  $\mathcal{S}_{\frac{1}{2}\sigma(j)}^{-1}$  is based on the observation that  $\mathcal{S}_{\frac{1}{2}\sigma(j)}$

TABLE I  
RMS SUBBAND INTENSITIES FOR SYNTHETIC  
IMAGE WITH SINGLE ORIENTATION

| Subband Label | Subband | Conventional | Resampled |
|---------------|---------|--------------|-----------|
| 1             | HH      | 5.78         | 1.92      |
| 1             | HL      | 2.15         | 2.00      |
| 1             | LH      | 6.58         | 8.51      |
| 2             | HH      | 18.42        | 1.19      |
| 2             | HL      | 6.95         | 2.49      |
| 2             | LH      | 20.50        | 27.60     |
| 3             | HH      | 55.21        | 0.87      |
| 3             | HL      | 20.69        | 1.46      |
| 3             | LH      | 60.49        | 81.66     |
| 3             | LL      | 1012.54      | 1011.82   |

is an approximation to an ideal shift by  $\frac{1}{2}\sigma(j)$  samples. It follows that  $\mathcal{S}_{-\frac{1}{2}\sigma(j)}$  is an approximation to  $\mathcal{S}_{\frac{1}{2}\sigma(j)}^{-1}$ .<sup>4</sup> Our proposed algorithm for computing  $\mathcal{S}_{\frac{1}{2}\sigma(j)}^{-1}\tilde{\varphi}$  for some sequence  $\tilde{\varphi}$  is to use iterative refinement in conjunction with this approximate inverse. In particular, from a sequence  $\tilde{\varphi}$ , we compute  $\varphi = \mathcal{S}_{\frac{1}{2}\sigma(j)}^{-1}\tilde{\varphi}$  as the limit of the iterative refinement sequence  $\varphi_0, \varphi_1, \varphi_2, \dots$ , where  $\varphi_0 = 0$  and

$$\varphi_k = \varphi_{k-1} + \mathcal{S}_{-\frac{1}{2}\sigma(j)}(\tilde{\varphi} - \mathcal{S}_{\frac{1}{2}\sigma(j)}\varphi_{k-1}). \quad (12)$$

The arguments above concerning the stability of inverting  $\mathcal{S}_{\sigma(j)}$  assumed that  $[K_1, K_2] = [-\infty, \infty]$ . Although, in practice,  $K_1$  and  $K_2$  are finite, we find empirically that our iterative refinement algorithm for computing  $\mathcal{S}_{\frac{1}{2}\sigma(j)}^{-1}$  converges rapidly when  $\frac{1}{2}\sigma(j)$  is not too close to  $\pm\frac{1}{2}$  and, in particular, when  $\frac{1}{2}\sigma(j) \in [-\frac{1}{4}, \frac{1}{4}]$ , the range to which  $\frac{1}{2}\sigma(j)$  is restricted by (5). Therefore, we see that the operator  $\mathcal{S}_{\pm\frac{1}{2}\sigma(j)}^{-1}$  exists for all  $j$  and can be implemented. Thus,  $\tilde{x}'$  can always be generated from  $x$  according to (11), and  $x$  may be readily recovered from  $\tilde{x}'$  using

$$x(l_1, \cdot) = \mathcal{S}_{\frac{1}{2}\sigma(l_1)}^{-1} \circ \mathcal{S}_{-\frac{1}{2}\sigma(l_1)}\tilde{x}'(l_1, \cdot). \quad (13)$$

It is important to note that when  $[K_1, K_2]$  approaches  $[-\infty, \infty]$ , the composed operator  $\mathcal{S}_{-\frac{1}{2}\sigma(j)}^{-1} \circ \mathcal{S}_{\frac{1}{2}\sigma(j)}$  represents LSI filtering with an all-pass frequency response. Thus, the energies of the sequences  $x$  and  $\tilde{x}'$  are identical. When finite length sequences are considered, this energy conservation is approximately obeyed.

We have argued that the two-stage interpolator of (11) is necessary in order to ensure that the resampling process is invertible. Even if we were prepared to sacrifice exact invertibility in order to use some simpler interpolator, we would, nevertheless, require our interpolator to possess this all-pass property. Otherwise, the interpolators for each line  $l_1$  of the image would, in general, all have different magnitude responses, which would severely impair the performance of the subsequent subband coder. The interpolator of (10), for example, which we rejected on the basis of invertibility, also has an insufficiently uniform magnitude response unless a very large number of taps are used.

<sup>4</sup>If  $\mathcal{S}_\sigma$  represented ideal sinc-interpolative shifting by  $\sigma$ , then we would have  $\mathcal{S}_\sigma^{-1} = \mathcal{S}_{-\sigma}$ .

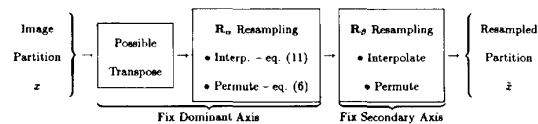


Fig. 3. Expanded schematic of resampling block in Fig. 1.

#### D. The Two Resampling Parameters

In the previous two sections, we have introduced a class of resampling matrices  $\mathbf{R}_\alpha$  and  $\mathbf{R}_\beta$  that satisfy constraints required for invertible resampling, and we have developed an algorithm for implementing the corresponding resampling operators. In our orientation adaptive subband scheme, the resampling applied to each partition is characterized by two parameters, corresponding to the orientations of the two resampling axes.

The resampling is accomplished in two phases. In the first phase, sampling of type  $\mathbf{R}_\alpha$  is applied. This corresponds to resampling axes, where one is oriented at an angle of  $\tan^{-1}(\alpha)$  to the vertical, whereas the other remains horizontal. We restrict

$$|\alpha| \leq 1 \quad (14)$$

which allows for a  $\pm 45^\circ$  distortion of the vertical axis. We also allow for the image to first be transposed. In this way, after the first resampling phase, the so-called 'dominant' resampling axis may take on any orientation at all, whereas the other resampling axis is aligned either horizontally or vertically, depending on whether the image was first transposed. Our first resampling parameter indicates the value of  $\alpha$  and whether or not the image has been transposed, thus providing the orientation of the dominant axis.

In the second resampling phase, a resampling of type  $\mathbf{R}_\beta$  is applied. This does not affect the orientation of the dominant axis but allows the other, so-called 'secondary,' resampling axis to depart from the horizontal or vertical. We restrict

$$|\beta| \leq 1 \quad (15)$$

in order to guarantee that the the index permutations described by (6), which were applied in the  $\mathbf{R}_\alpha$  and  $\mathbf{R}_\beta$  phases, do not conspire to violate the convexity of the image's region of support. This restriction also acts to prevent the orientations of the two resampling axes from approaching one another too closely. Our second resampling parameter, then, is the value of  $\beta$ . A schematic of the resampling process appears in Fig. 3.

#### E. Interpretation of Resampling

We conclude this section with a pictorial explanation to illustrate the significance of resamplings of type  $\mathbf{R}_\alpha$  and  $\mathbf{R}_\beta$  as well as the cascade  $\mathbf{R}_\alpha$  followed by  $\mathbf{R}_\beta$  from Fig. 3. Fig. 4(a) depicts a hypothetical image with square support containing a pair of oriented lines. The figure shows how a resampling of  $\mathbf{R}_\alpha$  with  $\alpha = \frac{1}{2}$  distorts the support of the image and realigns the oriented features in the vertical direction. Fig. 4(b) demonstrates the effect of a resampling of type  $\mathbf{R}_\beta$  on a similar hypothetical image. Finally, Fig. 4(c) illustrates the

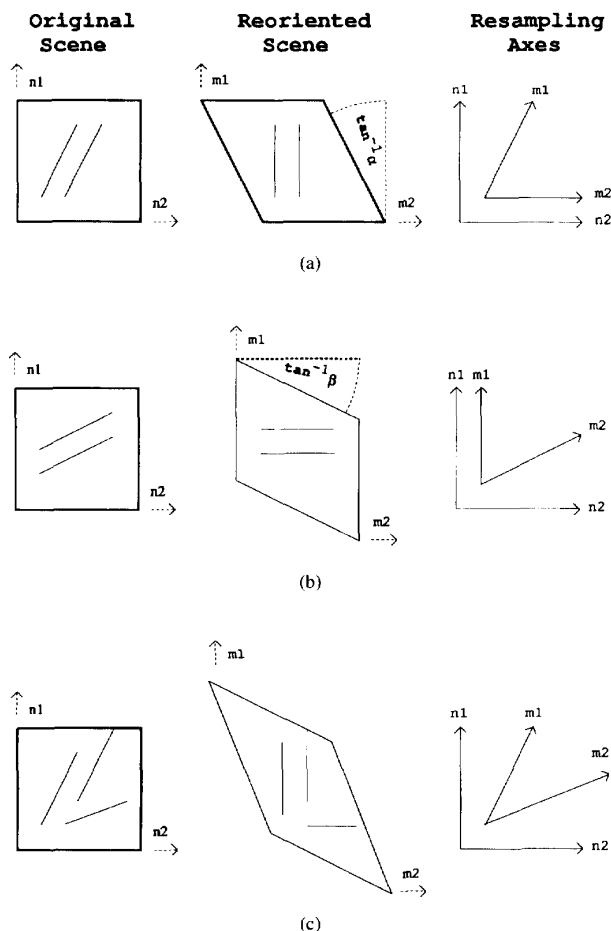


Fig. 4. Pictorial illustration of resamplings  $\mathbf{R}_\alpha$ ,  $\mathbf{R}_\beta$ , and  $\mathbf{R}_\alpha$  followed by  $\mathbf{R}_\beta$  in (a), (b), and (c) respectively.

ability of a combination of  $\mathbf{R}_\alpha$  and  $\mathbf{R}_\beta$  to realign features with two distinct orientations to lie in the vertical and horizontal directions. Naturally, the distortion of the region of support, due to resamplings  $\mathbf{R}_\alpha$  and  $\mathbf{R}_\beta$ , demands special consideration to be given to subband filtering and subsampling operations in the neighbourhood of the image boundaries. These issues are addressed in Section VI-A.

### III. ILLUSTRATIVE SYNTHETIC EXAMPLES

At this point, it is helpful to illustrate the performance of our orientation adaptive scheme with two synthetic images. To begin with, we consider a synthetic image of  $256 \times 256$  pixels consisting of uniform light and dark regions separated by a single oriented boundary. The boundary makes an angle of  $\tan^{-1}(0.75)$  with the vertical axis. After resampling the image in such a way as to realign this boundary to the vertical orientation, three levels of separable subband decomposition are applied. For comparison, we also apply the same separable subband decomposition directly to the synthetic image without any resampling. The root mean squared (RMS) intensities

of the subband samples in these two cases are compared in Table 1.

From this simple example, the effectiveness of our resampling strategy may be seen. When resampling is applied, the boundary information is compressed almost entirely into the low-frequency subband of the vertical subband filters—i.e., the ‘LL’ and ‘LH’ subbands. By contrast, the conventional scheme assigns as much energy to the ‘HH’ subbands as to the ‘LH’ subbands. This compaction of the signal information into fewer subbands allows our orientation adaptive subband scheme to achieve lower bit rates for a given choice of quantization parameters. The somewhat less impressive performance in the highest frequency subbands—i.e., the level 1 subbands—is due to the fact that the boundary does not quite form a perfect vertical line after resampling. This is a consequence of nonideality in the interpolator of (11). At low bit rates, however, the level 1 subbands usually account for little of the overall bit budget.

It is quite common in practice that a single image partition contains more than one orientational feature. To illustrate the effectiveness of our two parameter resampling scheme, which is discussed in Section II-D, in dealing with such cases, we consider the synthetic image of Fig. 5(a), which consists of two differently oriented boundaries between uniform light and dark regions. The image dimensions are again  $256 \times 256$  pixels. We compare the performance of a three-level separable subband decomposition applied to

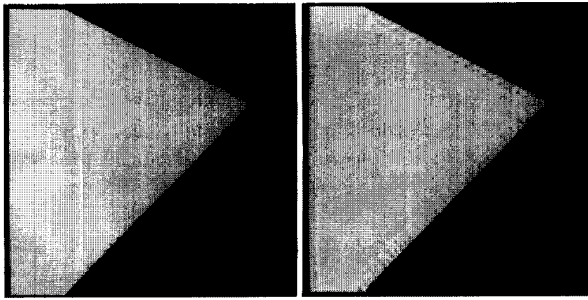
- the synthetic image itself
- the synthetic image resampled with only one of the two resampling stages discussed in Section II-D
- the synthetic image resampled with both resampling stages to align the two oriented boundaries both horizontally and vertically.

In this example, we observe the effectiveness of our resampling strategy in improving both the perceptual quality of the reconstructed image and the bit rate when quantization is coarse. In particular, the subbands in all three cases are quantized and coded using the strategies discussed in Section VI-B, with the quantization parameters appearing in the fourth line of Table II. The corresponding reconstructed images appear in Figs. 5(b)–(d). Gibbs phenomena at the boundary edges are clearly observed in Fig. 5(b), in which no resampling has been applied. In Fig. 5(c), the single resampling stage is able to orient one of the boundaries vertically. Consequently, this boundary is much more sharply rendered in the reconstructed image; however, the other boundary still displays the Gibbs phenomena. In Fig. 5(d), the boundary artifacts have been significantly reduced in both boundaries by the application of two resampling stages.

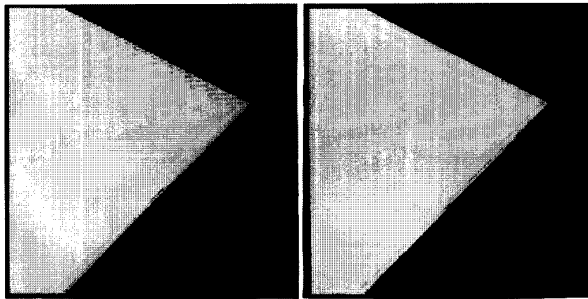
A similar pattern emerges in the bit rates corresponding to these three different subband decompositions. The bit rate corresponding to Fig. 5(b), in which no resampling is applied, is 0.071 b/pixel. The bit rate corresponding to Fig. 5(c) is 0.061 b/pixel. Finally, when both resampling stages are used, the bit rate corresponding to Fig. 5(d) is 0.055 b/pixel. We see, then, a progressive improvement both in perceptual quality of the reconstructed images and in bit rate

TABLE II  
QUANTIZATION PARAMETERS, BIT RATES, AND RECONSTRUCTION ERRORS FOR FIGS. 10 TO 16

| Coding Scheme | Figure | Bit Rate | PSNR (dB) | Quantization Parameters |                   |            |                     |
|---------------|--------|----------|-----------|-------------------------|-------------------|------------|---------------------|
|               |        |          |           | $\Delta$                | $\Gamma_{\Delta}$ | $\Upsilon$ | $\Gamma_{\Upsilon}$ |
| JPEG          | 10     | 0.199    | 28.2      | N/A                     |                   |            |                     |
| Conventional  | 11     | 0.199    | 30.8      | 91                      | 0.8               | 77         | 0.8                 |
| Adapted       | 12     | 0.199    | 30.5      | 90                      | 0.8               | 75         | 0.8                 |
| Conventional  | 13     | 0.098    | 28.2      | 220                     | 0.8               | 180        | 0.8                 |
| Adapted       | 14     | 0.097    | 27.9      | 240                     | 0.8               | 190        | 0.8                 |
| Conventional  | 15     | 0.302    | 32.4      | 65                      | 0.8               | 44         | 0.8                 |
| Adapted       | 16     | 0.303    | 32.2      | 65                      | 0.8               | 43         | 0.8                 |



(a)



(b)

Fig. 5. Two boundary synthetic image in (a) reconstructed from coarsely quantized subbands with (b) no resampling; (c) one resampling stage; (d) two resampling stages.

as the number of resampling stages is increased from none to two.

#### IV. ORIENTATION ADAPTATION

The choice of appropriate resampling axes is clearly at the heart of our algorithm. Our ambition here must be to locate the dominant linear features within each image partition, selecting the two resampling parameters discussed in Section II-D so that each of the resampling axes lies parallel to one or another of these features. It should be clear from the outset, however, that this is too vague an objective. In particular, it is not clear at which resolution we should search a partition for its linear features. Such a decision is a function of the target bit rate for the system. If we are interested in very low bit rates, then the linear features of interest will be those that appear at low resolutions, whereas at higher bit rates, we will be more interested in the features that appear at higher

resolutions. As a result of these considerations, we approach the choice of resampling axes in two distinct stages as outlined below. In the first stage, we examine each image partition at several different resolutions, collecting either two or four distinct pairs of resampling axes at each resolution. In the second stage, we decide which of these to apply to the final subband decomposition. As we shall see, this decision is based on an estimate of the coding performance associated with each pair of resampling axes.

##### A. Stage 1: Collection of Feasible Orientations

In this section, we provide details of our algorithm for generating a small collection of feasible resampling axis pairs for a given image partition. We begin in Section IV-A-1 by discussing the concept of image resolution and its application to our problem. In Section IV-A-2, we describe the simple edge detector applied to obtain edge information at each resolution. Finally, Section IV-D-3 describes the determination of dominant feature orientations from this edge information, whereby either two or four feasible pairs of resampling axes are assigned at each resolution.

1) *The Concept of Resolution:* The concept of resolution dependent feature extraction was introduced by Marr and Hildreth [14]. Following their work, we consider each of the image partitions, denoted by  $\mathcal{P}_i$ , at a succession of different resolutions derived by applying Gaussian low-pass filters of the form

$$g_r(n) = e^{-(n/\sigma_r)^2} \quad (16)$$

to the rows and columns of  $\mathcal{P}_i$  to obtain  $\mathcal{P}_i^r$ .

The significance of resolution to our proposed orientation adaptive subband scheme is that edge information typically appears in the high-frequency subbands of a subband system. Identification of the particular subbands influenced by a given edge segment is a resolution-dependent problem. A very smooth edge segment, for example, would be expected to have little influence on the highest frequency subbands and yet should appear in one or more of the lower frequency subbands. The significance of edges detected at a particular resolution depends on how coarsely the subbands, which respond to that resolution, are to be quantized.

It is for this reason that we choose the filter parameters  $\sigma_r$  in (16) so that the resolutions  $\mathcal{P}_i^r$  correspond roughly to the bandwidths encompassed by the different levels of the

subband coding system depicted in Fig. 1.<sup>5</sup> If a three-level subband decomposition is to be employed, for example, we consider each image partition at three resolutions:  $\mathcal{P}_i^0$  with full image bandwidth;  $\mathcal{P}_i^1$  with a bandwidth approximating that of the 'LL' band after one level of decomposition; and  $\mathcal{P}_i^2$  with a bandwidth approximating that of the 'LL' band after two levels of decomposition.

At each of these resolutions, a process of edge detection followed by orientation analysis is used to determine either two or four feasible pairs of resampling axes. The selection of the most appropriate pair of axes from the collection so obtained is the second stage in our orientation adaptation algorithm and is discussed in Section IV-B.

2) *Edge Detection*: To each partition  $i$  at each resolution  $r$ , we apply a relatively simple edge detection algorithm to derive a field of edges  $\mathcal{E}_i^r$  from  $\mathcal{P}_i^r$ . To do this, we first form the vertical and horizontal gradient fields

$$\mathcal{V}_i^r(n, m) = \mathcal{P}_i^r(n, m) - \mathcal{P}_i^r(n - 1, m)$$

and

$$\mathcal{H}_i^r(n, m) = \mathcal{P}_i^r(n, m) - \mathcal{P}_i^r(n, m - 1)$$

and then form the field of edges from the local maxima and minima of these gradient fields:

$$\begin{aligned} \mathcal{E}_i^r(n, m) &= \begin{cases} \mathcal{V}_i^r(n, m)^2 & \text{if } \{\mathcal{V}_i^r(n, m) - \mathcal{V}_i^r(n - 1, m)\} \cdot \\ & \{\mathcal{V}_i^r(n, m) - \mathcal{V}_i^r(n + 1, m)\} < 0 \\ 0 & \text{otherwise} \end{cases} \\ &+ \begin{cases} \mathcal{H}_i^r(n, m)^2 & \text{if } \{\mathcal{H}_i^r(n, m) - \mathcal{H}_i^r(n, m - 1)\} \cdot \\ & \{\mathcal{H}_i^r(n, m) - \mathcal{H}_i^r(n, m + 1)\} < 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

In this way, the nonzero pixels of  $\mathcal{E}_i^r$  are concentrated at the image edges that appear in the  $i$ th partition at the  $r$ th resolution. The amplitudes of the nonzero pixels  $\mathcal{E}_i^r$  are the squares of the corresponding edge gradients.

3) *Orientation Analysis*: The technique described in this section, for extracting orientation information from the fields of edges  $\mathcal{E}_i^r$  bears a strong resemblance to the Hough transform technique described by O'Gorman and Clowes [9] for linear edge enhancement. In our case, we are interested only in orientation information and not edge location. Furthermore, we are interested in determining orientation information in the form of the resampling parameters discussed in Section II-D with the constraints of (14) and (15).

We define an objective function  $\mathcal{O}_i^r(\theta)$  of orientation  $\theta$ , the local maxima of which correspond to dominant orientations in the field of edges  $\mathcal{E}_i^r$ . This objective function will be described shortly; however, we first describe the resampling axis pairs derived from  $\mathcal{O}_i^r$  for each partition  $i$  and resolution,  $r$ .

1. The first potential resampling axis pair has a dominant axis orientation corresponding to the maximum of  $\mathcal{O}_i^r$ . Its

<sup>5</sup>At this point, the reader may wonder why the subband analysis filters themselves are not selected for resolution analysis, rather than the Gaussian filters of (16). The subband analysis filters, however, may only be used to obtain the resolution  $\mathcal{P}_i^1$  without resorting to downsampling, an operation that would reduce the accuracy of our orientation analysis. Moreover, the Gaussian filters are considered to provide a more reasonable model of the human visual system and, hence, be more suited to the feature detection task at hand [14].

secondary axis orientation is chosen to be the largest local optimum of  $\mathcal{O}_i^r$  satisfying the constraints of (14) and (15), or if no such local optimum exists, the secondary axis is chosen to be perpendicular to the dominant axis.

2. A second potential pair of resampling axes is chosen to have the same dominant axis orientation as the first but with the secondary axis being either vertical or horizontal—i.e., the parameter  $\beta = 0$  in Fig. 3.
- 3&4 If a sufficiently large local optimum of  $\mathcal{O}_i^r$  exists, sufficiently removed from the optimum, a third and fourth pair of potential resampling axes are chosen by following the procedures for the first and second with the dominant axis orientation set to this local optimum.

Each of these axis combinations is empirically found to have potential in seeking the optimum partition resampling.

It is sufficient to consider axis orientations in the range  $-45^\circ \leq \theta < 135^\circ$  relative to the vertical. Observing that when  $45^\circ < \theta < 135^\circ$ ,  $\mathcal{O}_i^r(\theta)$  may be computed by transposing  $\mathcal{E}_i^r$  and making the substitution  $\theta \leftarrow 90^\circ - \theta$ , we consider only the case  $|\theta| \leq 45^\circ$ . Our objective function  $\mathcal{O}_i^r(\theta)$  is obtained as follows.

We first generate  $\mathcal{E}_i^r(\theta)$  by reorienting  $\mathcal{E}_i^r$  with a resampling of  $\mathbf{R}_\alpha$  as defined in Section II-B, with  $\alpha = \tan^{-1} \theta$ . The interpolation phase represented by (5) is too computationally costly for evaluation of our objective function and is therefore omitted. This leads to a spatial distortion in the reoriented field of edges  $\mathcal{E}_i^r(\theta)$  of no more than  $\pm \frac{1}{2}$  a pixel. Our objective function is chosen not to be significantly affected by this distortion. Thus,  $\mathcal{E}_i^r(\theta)$  is obtained from  $\mathcal{E}_i^r$  by simple index permutation represented by (6).  $\mathcal{O}_i^r(\theta)$  is a measure of the vertical alignment of the edges appearing in  $\mathcal{E}_i^r(\theta)$ . In particular, letting  $l_k(\theta)$  denote the  $k$ th column of  $\mathcal{E}_i^r(\theta)$ ,  $\mathcal{L}_{l_k(\theta)}$ , the length of the  $k$ th column, and  $\mathcal{G}_{l_k(\theta)}$  the pixel sum along the  $k$ th column, we define

$$\mathcal{O}_i^r(\theta) = \sum_k \left( \frac{\mathcal{G}_{l_k(\theta)}}{\mathcal{L}_{l_k(\theta)}} + \frac{\mathcal{G}_{l_{k+1}(\theta)}}{\mathcal{L}_{l_{k+1}(\theta)}} \right)^2 \cdot (\mathcal{L}_{l_k(\theta)} + \mathcal{L}_{l_{k+1}(\theta)}). \quad (17)$$

The significance of this objective function will now be explained. We first consider the simpler objective function

$$\hat{\mathcal{O}}_i^r(\theta) = \sum_k \underbrace{\frac{\mathcal{G}_{l_k(\theta)}}{\mathcal{L}_{l_k(\theta)}}}_1 \cdot \underbrace{\frac{\mathcal{G}_{l_k(\theta)}}{\mathcal{L}_{l_k(\theta)}} \cdot \mathcal{L}_{l_k(\theta)}}_2. \quad (18)$$

We may think of the term

$$\frac{\mathcal{G}_{l_k(\theta)}}{\mathcal{L}_{l_k(\theta)}} \quad (19)$$

as the density of gradient values contained in the line  $l_k$ . We now observe that if term 1 in (18) were omitted, we would have  $\hat{\mathcal{O}}_i^r(\theta) = \sum_k \mathcal{G}_{l_k(\theta)}$ , which is independent of  $\theta$ . The same observation would hold if the gradient density terms of (19) were independent of  $k$ . When the gradient density terms vary with  $k$ , however, larger values of term 2 in (18) are magnified by the presence of term 1, whereas smaller values are diminished. In this way, (18) is a measure of the variability in the gradient density terms of (19) with  $k$ .



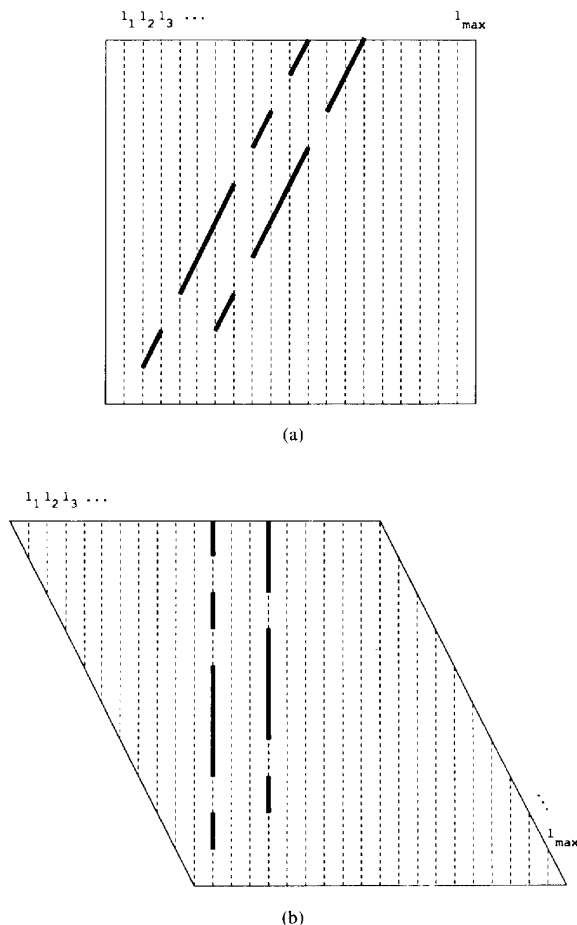


Fig. 6. Explanation of orientation analysis objective function  $\mathcal{O}'_i(\theta)$ : (a)  $\mathcal{E}'_i(0)$ ; (b)  $\mathcal{E}'_i(\theta)$ ,  $\tan(\theta) = \frac{1}{2}$ .

If the edges in  $\mathcal{E}'_i(\theta)$  do possess some discernible orientation, then the gradient density terms of (19) will vary most with  $k$  when this orientation is vertical, thus maximizing  $\hat{\mathcal{O}}'_i(\theta)$  when  $\theta$  corresponds to the dominant orientation of edges in  $\mathcal{E}'_i$ . Consider the example of Fig. 6 in which  $\mathcal{E}'_i$  contains two similarly oriented features. In Fig. 6(b), corresponding to  $\theta = \tan^{-1}(\frac{1}{2})$ , the gradient density terms are concentrated into just two of the lines  $l_k$ , whereas in Fig. 6(a), corresponding to  $\theta = 0$ , the gradient density terms are much more uniform.

As a further test of the reasonableness of the objective function of (18), suppose that  $\mathcal{E}'_i$  contains a random distribution of edge intensities, in which case, no discernible orientational information is present. Then, the gradient density terms of (19) would be roughly independent of  $k$ , and therefore,  $\hat{\mathcal{O}}'_i(\theta)$  would be independent of  $\theta$ , as expected.

Our objective function in (17) has the same properties as that of (18), except that it measures the variability with  $k$  of sums of gradient densities from adjacent lines. This modification is found to make our objective function less sensitive to noise in the field of edges  $\mathcal{E}'_i$ .

Another important feature of our objective function is that it responds to broken as well as to unbroken edge segments.

This is because it is formulated in terms of the sums  $\mathcal{G}_{l_k(\theta)}$ , whose values are little affected by edge continuity. It shares this property with the Hough transform [9], which has been applied to the completion of broken edge segments. This lack of dependence on edge continuity allows us to skip the interpolation phase in reorienting  $\mathcal{E}'_i$  to form  $\mathcal{E}'_i(\theta)$ , as described above.

In developing this objective function, we considered a number of alternative approaches, in which the orientations of dominant edge segments, selected according to their length, were used as a measure of dominant orientation. These approaches were found to suffer from their restricted view of orientation as that of an isolated edge segment. In contrast, the objective function of (17) is a function of every pixel in  $\mathcal{E}'_i$ , making it effective in assigning appropriate significance to collections of similarly oriented edge segments.

### B. Stage 2: Selection from Feasible Orientations

As already mentioned, the resolution at which dominant linear features should be matched by the resampling axes depends on the target bit rate of our system. Based on this observation, we make our selection from amongst all of the feasible pairs of oriented axes, which are described in Section IV-A, based on an estimate of the corresponding bit rate. In addition, we consider the bit rate associated with the conventional vertical/horizontal axis pair. The resampling axis pair giving the lowest bit rate is selected.

This strategy may be justified by the following two arguments. First, the examples presented in Section III demonstrate the compaction of signal energy into few subbands when the resampling axes correspond to dominant orientations. When quantization is coarse, the majority of the bit budget is assigned to the lower frequency subbands, and the code rate is most sensitive to the ability of a given pair of resampling axes to compact low-resolution features into as few as possible of these subbands. On the other hand, when quantization is fine, the majority of the bit budget is assigned to the higher frequency subbands<sup>6</sup> so the code rate is most sensitive to the alignment of resampling axes with the features appearing at higher resolutions. The bit rate is thus an appropriate method for comparing the significance of the directional features located at different resolutions in Stage 1. When quantization is coarse, the strategy assigns greater significance to the low-resolution features. These are also the features that have greatest perceptual significance when quantization is coarse. When quantization is fine, it assigns greater significance to the high-resolution features. These, again, are the features that have greatest perceptual significance when quantization is fine.

A second justification for the use of bit rate in selecting the resampling axes is that it prevents increases in the overall bit rate above that of conventional subband coding, which might otherwise offset the enhancement in perceptual edge quality that orientation adaptivity offers. In practice, we find that our algorithm seldom selects conventional resampling axes for the

<sup>6</sup>Even though fewer bits per pixel are assigned to the higher frequency subbands, they contain many more pixels than the lower frequency subbands.



Fig. 7. Original  $512 \times 512$  Lena test image.



Fig. 8.  $512 \times 512$  Lena reconstructed from orientation-adapted subbands at 0.2 bpp before smoothing of partition boundaries.

natural images we have considered, indicating the prevalence of useful orientational information.

The estimated bit rate used in this selection procedure involves the same quantization and run-length coding techniques to be described in Section VI-B; however, the full Huffman coding of Section VI-B is too computationally expensive to be used at this point where only an estimate of the bit rate is required. Instead, our estimate is obtained from an approximate entropy figure based on statistical information from other images.

#### V. REMOVING ARTIFACTS AT PARTITION BOUNDARIES

The algorithm we have divulged so far treats each partition of the image entirely independently, leading to blocking artifacts in the reconstructed image when quantization is coarse. The task of removing these is somewhat simplified by the fact that our partitions are chosen to be relatively large— $64 \times 64$  pixels usually—and the artifacts are of known location and observed to be of small extent. Fig. 8 demonstrates the nature of these blocking artifacts. Removal of the artifacts is performed in the receiver after image reconstruction.

With the objective of perturbing the image as little as possible, our approach is to smooth the partition boundaries by means of a projection operator. A simple and efficient means to do this is to take advantage of the orthogonality of our subband filter system, which is discussed in Section VI-A. The idea is to assert that the high-frequency subbands of a 1-D subband decomposition taken across the partition boundary be zero in the neighborhood of the boundary. In Section V-A, we develop this idea for smoothing in the direction perpendicular to the partition boundaries. In Section V-B, we describe our more general, oriented smoothing algorithm, by which we further exploit directional information in the image.

#### A. Elementary Approach

Consider a vertical partition boundary, and let the  $k$ th image row intersecting this boundary be denoted  $\lambda_k$ . Each such row forms a 1-D sequence of data  $\lambda_k(n)$  with  $n = 0$  and  $n = 1$  arbitrarily chosen to correspond to boundary pixels. Next, we decompose each of these sequences into subbands. To do this, we apply a two-channel subband decomposition using the same subband filters  $\mathbf{F}$  and  $\mathbf{G}$  of Fig. 1 iterating along the low-pass channel for  $L$  levels.<sup>7</sup> In this way, we generate  $L$  high-pass subbands with octave related passband widths and one low-pass subband with nominal bandwidth of  $\pi/2^L$ . Denoting these subband samples by  $\mathcal{H}_k^1(n), \dots, \mathcal{H}_k^L(n)$  and  $\mathcal{L}_k^L(n)$  respectively, we have

$$\begin{aligned} \lambda_k(n) = & \sum_{p_1} (g(n - 2p_1) \cdot \mathcal{H}_k^1(p_1) + f(n - 2p_1) \\ & \times \sum_{p_2} (g(p_1 - 2p_2) \cdot \mathcal{H}_k^2(p_2) + f(p_1 - 2p_2) \\ & \times \dots \\ & \times \sum_{p_L} \left( g(p_{L-1} - 2p_L) \cdot \mathcal{H}_k^L(p_L) \right. \\ & \left. + f(p_{L-1} - 2p_L) \cdot \mathcal{L}_k^L(p_L) \right) \dots) \end{aligned} \quad (20)$$

where the low-pass subband filter  $f(n)$  is symmetric about  $n = 0$ , and the high-pass subband filter  $g(n)$  is symmetric about  $n = 1$ .

For each subband level  $l \leq L$ , we set  $\mathcal{H}_k^l(p) = 0$  for  $-\Phi_l \leq p < \Phi_l$ . We then reconstruct  $\lambda'_k$  from these modified subbands according to (20), replacing the original  $k$ th row  $\lambda_k$  with  $\lambda'_k$ . Turning our attention to the constants  $\Phi_l$ , we observe from (20) that the zeroed subband samples at level  $l$  affect only the

<sup>7</sup>The value of  $L$  is a parameter in our algorithm, which we have found experimentally to be best set to  $L = 2$ .

samples  $\lambda'_k(n)$  with  $|n| \leq N_l$ , where

$$N_l = 2^{l-1}(2\Phi_l - 1) + M(2^l - 1),$$

in which the subband filters have  $2M + 1$  taps. We choose a small value for  $\Phi_L$  and then, for each  $l < L$ ,  $\Phi_l$  is chosen to be the largest integer such that  $N_l \leq N_L$ . In experimental work, for example, we have found  $\Phi_L = 2$  to give good results.

In this way, our algorithm is asserting that in the vicinity of the partition boundary, the modified lines  $\lambda'_k$  have spectral bandwidth of no greater than approximately  $\pi/2^L$ . By orthogonality of the subband system, the operation just described is also a projection operator.

A similar algorithm is applied to the horizontal partition boundaries in which the lines  $\lambda_k$  are image columns that intersect the boundary.

It should be noted that the algorithm steps outlined above are only didactic. We need not generate the complete subband decomposition of each line  $\lambda_k$  but only those subband samples that are to be set to zero. Synthesizing from these samples alone and subtracting the result from the original line has the same effect and is numerically efficient.

### B. Oriented Smoothing at Partition Boundaries

The method described above works well in removing partition boundary artifacts from smooth regions of an image; however, it is not uncommon for oriented edges in the image to intersect with partition boundaries. The elementary approach generates objectionable artifacts at such edges. An example of these artifacts appears in Fig. 18, which is discussed in Section VII. In order to minimize such artifacts, we choose the lines  $\lambda_k$ , which are described in Section V-A, to be oriented approximately parallel to any edges which cross a given partition boundary.

In particular, for each of the four boundaries of each partition, we determine an orientation  $\theta$  for the smoothing. Pixels in the neighborhood of each boundary are assigned to oriented lines  $\lambda_k$  using the same index permutation method described in Section IV-A-3, thereby allowing for orientations between  $\pm 45^\circ$  relative to the normal direction to the boundary. The subband-based smoothing of these lines then proceeds exactly as in Section V-A. The appropriate subband samples are denoted by  $\mathcal{H}_{k,\theta}^1(n), \dots, \mathcal{H}_{k,\theta}^L(n)$  and  $\mathcal{L}_{k,\theta}^L(n)$ .

We determine the optimum orientation for this smoothing by minimizing  $\sum_k \sum_{p=-\Phi_L}^{\Phi_L-1} \mathcal{H}_{k,\theta}^L(p)^2$  as a function of  $\theta$ , where the summation is over all lines  $\lambda_k$  crossing the boundary under consideration. In practice, the visual effect is not highly sensitive to the orientation chosen, and therefore, only 16 different orientations  $\theta$  are considered, and overall image reconstruction time is not significantly compromised. An analysis of the complexity issues associated with the smoothing algorithms discussed in this section appears in Section VII.

## VI. EXPERIMENTAL PARAMETERS AND RESULTS

The object of this section is to present a comparison of reconstructed images using conventional and orientation adaptive subband coding approaches. This comparison appears in Section VI-D; however, we first present details of a number

of the choices made in building our subband system. In Section VI-A, we give details of the partitioning of images into smaller regions and the filtering considerations relevant to subband decomposition of these small regions. In Section VI-B, we describe the quantization and coding procedures applied to subband samples and auxiliary information in order to obtain bit rates for comparison with other image coding techniques. In Section VI-C, we indicate the parameters chosen for removal of partition boundary artifacts.

### A. Filters, Partitioning, and Sample Rate Preservation

Our approach to oriented subband coding via the resampling process described in Section II provides two degrees of orientational freedom, namely, the two resampling axes, for each image partition. Although finer partitioning increases our ability to exploit the orientational information in the image, it does not come without cost. An obvious cost is the appearance of blocking artifacts described in Section V. A second obvious cost is the increase in orientation information that must be transmitted as the number of partitions increases. A third cost is associated with degradation in the subband system's performance as the partitions decrease in size. This is due to the increase in the proportion of pixels that lie close to a partition boundary. We choose to use the well-known symmetric-extension method [21] to obtain exact sample rate preservation in our subband coder; however, in the vicinity of partition boundaries, the effective subband filter responses suffer from this symmetric extension.

The symmetric-extension technique requires linear phase filters, whereas we would also like to take advantage of the quantization noise transmission properties of orthogonal subband systems [5], [22]. Linear phase and orthogonality are known to be mutually exclusive requirements in perfect reconstruction subband systems [23], and therefore, we choose to use the 9 tap almost perfect reconstruction, almost orthogonal subband filters of Adelson *et al.* [1]. The reconstruction error with these filters is practically never as large as 1% with three levels of subband decomposition and is never observable.

We choose to employ a three-level subband decomposition, for which we find a good compromise partition size to be  $64 \times 64$  pixels. This means that the lowest level subbands are generated by applying the  $9 \times 9$  tap separable subband filters to a decimated image consisting of 256 pixels with the parallelogram region of support discussed in Section II. Recalling that a separable subband filter system is implemented by applying 1-D subband decompositions to the rows and the columns of the resampled image, it is apparent that when this region of support is not rectangular, the rows or columns near two of its vertices will become too short to use the symmetric-extension method. For these regions, we resort to the use of the simplest known subband filters having the two tap impulse responses

$$\left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \quad \text{and} \quad \left( \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right).$$

The parameter  $N_{\max}$  of (9) is chosen to be  $N_{\max} = 10$ , giving a maximum sinc interpolation filter size of 21 taps. It is not yet clear how small  $N_{\max}$  can be before

coding performance is compromised. The iterative refinement algorithm given in (12) is allowed three iterations, in which it typically converges to within an error of less than 0.1% of the range of the image pixel values.

### B. Quantization and Coding

In order to demonstrate the efficacy of our orientation adaptive subband scheme as well as to indicate the regime of bit rates where its application might be most beneficial, it is necessary to code our subbands. We adopt a fairly traditional approach in which the lowest frequency subband is DPCM coded and the others run-length coded. The quantizers are uniform with a dead zone about zero, except in the case of the DPCM subband, which has a straightforward uniform quantizer. Each subband of each partition is scanned separately using either a horizontal, a vertical, a top left to bottom right diagonal, or a top right to bottom left diagonal scheme. Finally, the resulting streams are Huffman coded. Separate Huffman codebooks are maintained for each subband, and within the run-length coded subbands, separate Huffman codebooks are kept for the black run lengths, the white run lengths, and the nonzero pixel values.

1) *Quantization*: It is a fundamental assumption in subband coding that the quantization of different subbands may be optimized to match the spectral sensitivity of the human visual system. In order to accommodate such adjustments, our quantizers are specified by four parameters: a base quantization interval  $\Delta$ , a base dead zone threshold  $\Upsilon$ , a scaling coefficient for the quantization interval  $\Gamma_\Delta$ , and a scaling coefficient for the dead zone threshold  $\Gamma_\Upsilon$ . In any particular subband, the quantization interval is  $\Delta(\Gamma_\Delta)^k$ , where  $k$  is the number of low-pass filtering stages that have been applied in arriving at the subband. Similarly, the dead zone threshold is  $\Upsilon(\Gamma_\Upsilon)^k$ . Thus, for example, at the first subband decomposition level, the HH subband would correspond to  $k = 0$ , and the HL and LH subbands to  $k = 1$ . At the next level, the HH subband would correspond to  $k = 2$ , and so on.

2) *Auxiliary Information*: In addition to the subbands themselves, there is additional information that must be transmitted in our adaptive scheme, foremost of which is the axis orientation information. Recalling that the resampling parameters  $\alpha$  and  $\beta$  are constrained to lie between  $\pm 1$ , we uniformly quantize them each to 128 levels. An additional bit is required to signal transposition in the resampling schematic of Fig. 3. Due, however, to the relatively common occurrence that a partition's axes be merely the conventional vertical and horizontal directions, a leading bit is allocated to signify whether or not this is the case. Finally, an extra bit is allocated to signify whether the secondary axis is in one of the vertical or horizontal directions—i.e.,  $\beta = 0$ , which is also a common occurrence. In all then, the coding of each partition's resampling information requires 1, 10, or 17 bits.

An additional 2 bits are required for every subband of every partition to encode the scanning method used—horizontal, vertical, or one of two diagonals. The code rate corresponding to each of these scanning methods is compared before transmission. The only event in which these 2 bits need not be

transmitted is when no subband data is to be transmitted; this is an event signaled by the fact that the run-length codes for that subband contain only the end-of-run symbol.

### C. Partition Boundary Artifacts

Fig. 8 shows the effect of reconstructing the standard test image (Lena) from orientation adapted subbands at 0.2 bits per pixel before any attempt is made to remove partition boundary artifacts. Fig. 12 shows the same image after removal of these artifacts using the algorithm presented in Section V. We choose  $L = 2$  and  $\Phi_L = 2$ . With our 9 tap filters, this leads to a value of  $\Phi_1 = 5$  when applying the procedure outlined in Section V-A. We find that  $L = 1$  provides insufficient smoothing, whereas  $L = 3$  produces its own noticeable artifacts in the image. These choices work well on all of the other natural images and bit rates we have considered.

### D. Comparison of Reconstruction Qualities

The results presented in this section are all related to the standard test image (Lena); however, we apply our scheme to a total of ten natural images, the statistics from which are used to build our Huffman codebooks in order to obtain reliable bit rates. The Lena image is shown in Fig. 7. We compare the quality of reconstructed versions of the Lena image obtained using three different coding techniques at equivalent bit rates: the JPEG image compression standard, conventional separable subband coding, and our orientation adaptive subband scheme.

A conventional separable subband scheme is obtained by applying our orientation adaptive subband coding software in the case where the partition size is that of the whole image—rather than the  $64 \times 64$  partition size used in the adaptive case—and forcing the resampling axes to always be horizontal and vertical. In this way, exactly the same subband filters and quantization and coding strategy are used for the conventional and adaptive subband schemes, allowing a meaningful comparison of performance between the two approaches. The Huffman codebooks for adaptive and conventional subband approaches are maintained separately, where each is derived from subband statistics obtained from the ten natural images, one of which is the Lena image itself. The quantization parameters discussed in Section VI-B-1 are obtained by experimentation. In this process, the parameters are varied in such a way as to maintain a given bit rate while the quality of the reconstructed image is subjectively assessed.

In applying JPEG, we use nonuniform quantization tables that have already been optimized for natural images. To provide a fair comparison with the subband coding schemes, we build the Huffman codebooks for JPEG from the statistics of the same ten images used to build our subband system codebooks. The JPEG bit rate is based on the assumption that neither quantization tables, Huffman tables, nor resynchronization codes are to be transmitted—i.e., an identical communication scenario to that for the subband coders. In addition, the reconstructed JPEG image is smoothed with a Gaussian low-pass filter in order to reduce objectionable blocking artifacts. Such post-processing is common practice for low bit rate JPEG images.

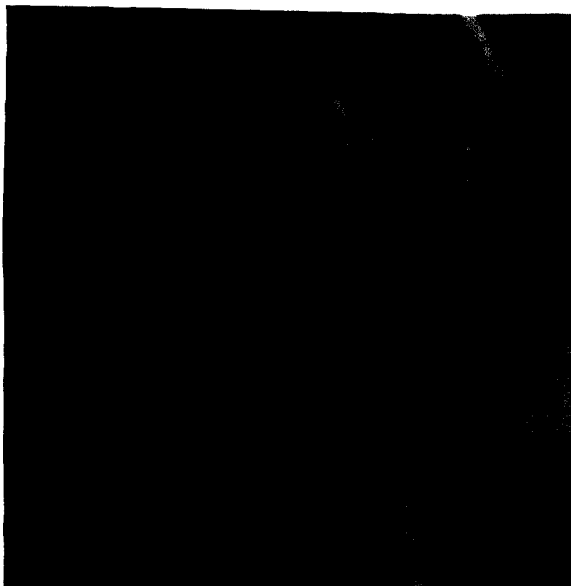


Fig. 9. Resampling axes for  $512 \times 512$  Lena with  $64 \times 64$  partitions.



Fig. 11.  $512 \times 512$  Lena reconstructed from conventional separable subbands at 0.2 bpp.



Fig. 10.  $512 \times 512$  Lena reconstructed from JPEG bitstream at 0.2 bpp.



Fig. 12.  $512 \times 512$  Lena reconstructed from orientationadapted subbands at 0.2 bpp.

A comparative study of reconstructed images at various low bit rates appears in Figs. 10 through 15, with the reconstructed image properties appearing in Table II. Figs. 11 and 12 compare reconstructions of the Lena image using orientation adaptive and conventional subband systems at 0.2 bits per pixel. At this bit rate, the reconstructed image is of moderate quality with the Gibbs phenomena at image edges in Fig. 11 being arguably the most objectionable artifacts, especially in the regions of the hat, the shoulder, and the mirror edges. These artifacts are all but absent in Fig. 12, demonstrating the success of our orientation adapted approach. It is interesting that both subband schemes appear to give higher quality reconstructed

images than that for JPEG, which is shown in Fig. 10. In Fig. 9, the resampling axis orientations, which are obtained in the process of generating Fig. 12 at 0.2 bits per pixel, are displayed in red for the dominant axis and green for the secondary axis within each of the image partitions. Note that our algorithm is highly successful in locating the dominant linear features within the image, often exploiting the availability of both axis orientations to capture as much directional information as possible.

We also show reconstruction results at 0.1 bits per pixel



Fig. 13.  $512 \times 512$  Lena reconstructed from conventional separable subbands at 0.1 bpp.



Fig. 15. Corner of  $512 \times 512$  Lena reconstructed from conventional separable subbands at 0.3 bpp.



Fig. 14.  $512 \times 512$  Lena reconstructed from orientation-adapted subbands at 0.1 bpp.



Fig. 16. Corner of  $512 \times 512$  Lena reconstructed from orientation-adapted subbands at 0.3 bpp.

in Figs. 13 and 14. At this bit rate, the image quality has suffered significantly; however, the Gibbs phenomena around image edges, which our algorithm removes, are still arguably the most objectionable artifacts in Fig. 13.

Finally, the performance of our algorithm at 0.3 bits per pixel is examined in Figs. 16 and 15. The Gibbs phenomena at  $512 \times 512$  Lena image edges are no longer as noticeable at this bit rate; therefore, the advantages of our orientation adaptive scheme over the conventional separable subband coder are

not as noticeable. In order to reveal the effect, Figs. 15 and 16 show an expanded view of the top right-hand corner of the Lena image.

Table II presents our quantization parameter choices as well as the bit rates and peak signal to noise ratios (PSNR's)<sup>8</sup> for each of these reconstructed images. We may make a number of observations concerning the information presented

<sup>8</sup>PSNR  $\triangleq 10 \log_{10} \frac{255^2}{\text{MSE}}$ , where MSE is the mean squared error.

TABLE III  
ARITHMETIC COMPLEXITY OF ORIENTATION ADAPTIVE SUBBAND RECEIVER

| Component                      | Floating Point Operations per Pixel                                 |  |
|--------------------------------|---|--|
|                                | Functional Form   | $M = 4, N_{\max} = 10,$<br>$P_x = P_y = 64, J = 3$ |
| Separable Subband Filtering    | $2\frac{2}{3}(4M+1)$  | 45   |
| Interpolation                  | $4(4N_{\max}+1)(J+1)$   | 656  |
| Smoothing of Section V-A<br>OR | $(\frac{1}{P_x} + \frac{1}{P_y})(2 \times 19 + 2 \times 10)(4M+1)$  | 31   |
| Smoothing of Section V-B       | $(\frac{1}{P_x} + \frac{1}{P_y})(17 \times 19 + 2 \times 10)(4M+1)$ | 182  |

in Table II. First, we draw attention to the  $\Gamma_{\Delta}$  and  $\Gamma_{\gamma}$  parameters. Subjective experimentation with these parameters has led to the conclusion that they should both assume a value of approximately 0.8. This means that the ratio between the quantization interval of the lowest frequency subband and that of the highest frequency subband should be about  $\frac{1}{4}$ , which is in agreement with the thesis that lower spatial frequencies have a greater perceptual significance than higher frequencies.

Although the orientation adaptation process itself always results in some coding gain, as discussed in Section IV-B, the communication overhead associated with each partition offsets this advantage. It is for this reason that at the lowest bit rate, quantization is slightly less coarse for the conventional approach than for our orientation-adapted approach. At higher bit rates, the adaptive scheme regains its footing. Also of interest from Table II is the fact that the PSNR is always slightly higher for conventional subband coding than for our adaptive approach. This is easily understood as the result of the image partitioning process that compromises the orthogonality of the subband system and hence the quantization noise energy near partition boundaries.

## VII. COMPLEXITY

The purpose of this section is to provide an indication of the numerical complexity of our orientation adaptive subband scheme for the purpose of comparison with conventional separable subband coding. In particular, we show that the receiver need only be approximately twice as numerically intensive as a separable subband receiver, whereas the transmitter is a couple of orders of magnitude more complex. The orientation adaptive scheme is thus highly asymmetric, unlike conventional separable subband coding. In Section VII-A, we discuss the influence of parameter choices on the receiver computational complexity as well as on quality of the reconstructed images. In Section VII-B, we indicate the approximate complexity of the transmitter.

### A. Receiver Numerical Complexity

Considering the computational elements that the receiver must embody, the following are of significance in terms of arithmetic operations: conventional separable subband filtering, the interpolation identified by (13); and the partition boundary smoothing described in Section V. We proceed to analyze the arithmetic complexity of each of these three elements. To do this, we assume that the image is divided into

partitions containing  $P_x \times P_y$  pixels, that the subband filters have  $2M+1$  taps, and that the number of iterations of the algorithm described by (12) is  $J$ . We assume for simplicity that the number of subband levels is infinite as only the first few subband levels significantly affect the numerical complexity. Also for simplicity, we assume that the partition boundary smoothing parameter choices described in Section VI-C are to be used. In particular, we have  $L=2, \Phi_2=2$  and  $\Phi_1=5$ .

The arithmetic complexities of each of the receiver's computational elements appear in Table III. The table also gives the complexities corresponding to the parameter choices of Section VI—i.e.,  $M=4, N_{\max}=10, P_x=P_y=64$ , and  $J=3$ . The images reconstructed from orientation-adapted subbands in Section VI-D were smoothed with the orientation sensitive smoothing algorithm discussed in Section V-B. With these choices, interpolation clearly dominates the receiver's arithmetic complexity. In all, the orientation adaptive receiver is about **20** times more numerically intensive than the receiver for a conventional separable subband system.

We may observe, however, that the interpolation phase corresponds to fractional shifting of image rows or columns by no more than  $\pm\frac{1}{2}$  a pixel, where the majority of the resampling task is accomplished by index permutation. If the interpolation phase is omitted entirely from the receiver, we would expect to observe slight staircase effects around oriented edges within the image. This is exactly what is observed in Fig. 17. Fig. 18, however, demonstrates a more satisfactory compromise between reconstruction quality and numerical complexity. In this case, we have chosen to employ a crude interpolative filter in the receiver with  $N_{\max}=2$  and no iterative refinement—i.e.,  $J=0$ . In addition, we have employed the simple boundary smoothing approach of Section V-A. The arithmetic complexity in this case is only **2.5** times that for a conventional separable subband system's receiver, and the quality of the image in Fig. 18 is arguably comparable with that of Fig. 12. The most noticeable new artifacts in Fig. 18 are those introduced by partition boundary smoothing where oriented edges intersect partition boundaries. Such artifacts may, for example, be seen along the rim of the hat and in the shoulder. They are practically absent in Fig. 12, in which the more sophisticated, orientation-sensitive algorithm of Section V-B is employed.

We conclude that at low bit rates, a simpler receiver with approximately twice the arithmetic complexity of a conventional subband receiver is probably quite sufficient. We also observe that almost all processing is performed in isolated



Fig. 17.  $512 \times 512$  Lena reconstructed from orientation-adapted subbands at 0.2 bpp without interpolation.



Fig. 18.  $512 \times 512$  Lena reconstructed from orientation-adapted subbands at 0.2 bpp with crude interpolation.

image partitions. This introduces regularity into our subband system, which may be exploited in a hardware implementation.

### B. Transmitter Numerical Complexity

It is evident that the transmitter associated with our orientation-adaptive scheme must be considerably more complex than the receiver. This is because the adaptation algorithm described in Section IV requires that the analysis system of Fig. 1 be executed once for each of a collection of

potential resampling axis pairs. As discussed in Section IV-A-3, there may be up to four such potential resampling axis pairs for every level in the subband filter pyramid. Corresponding to the parameter choices of Section VI, up to 12 resampling axis pairs must be checked, with each check requiring  $45+656$  floating point operations per image pixel. These figures are taken from Table III due to the symmetry of subband analysis and synthesis and of interpolation and inverse interpolation given by (11) and (13).

In addition, for each level in the subband filtering pyramid and for each of the 128 quantized angles of Section VI-B-2, we must evaluate the orientation objective function of (17), which involves approximately one floating-point operation per pixel being dominated by the evaluation of the  $\mathcal{G}_{l_k}(\theta)$  terms. We ignore the more modest requirements of the resolution reduction and edge detection phases of Sections IV-A-1 and 2.

In all then, for three levels of subband filtering, our proposed orientation-adaptive subband transmitter requires approximately  $12 \times (45 + 656) + 3 \times 128 \approx 9000$  floating-point operations per image pixel. This is about 200 times the computational power required by a conventional separable subband transmitter and about 100 times the computational power of the simple orientation adaptive receiver discussed in Section VII-A and used to generate Fig. 18. We note however that all processing in the transmitter takes place in isolated image partitions, introducing the possibility of highly parallel implementations.

## VIII. CONCLUSIONS

The work presented in this paper demonstrates the value in adapting a subband system to the local orientational features of an image. At low bit rates, we are able to achieve considerably cleaner edges than conventional approaches in which orientation is not considered. For natural images, Table II reveals that this subjective improvement in edge quality is not associated with significant changes in quantization precision but arises instead from directional attributes of the quantization noise. In the light of such results, we consider this directional shaping of quantization noise to be the most compelling motivation for orientation adaptive subband coding of images.

The particular orientation adaptive approach outlined here is capable of exploiting up to two distinctly oriented features in each region of an image and may be implemented using only the operations of 1-D filtering and index permutation. The transmitter is necessarily much more complex than that for a conventional separable subband system; however, a simplified form of the receiver exists, the numerical complexity of which is not substantially higher than that of the conventional subband receiver.

One could conceivably propose an alternative orientation-adaptive subband scheme in which directional sensitivity is achieved by a 2-D, nonseparable filter bank. For such an approach to be successful, however, a number of serious difficulties would need to be overcome. One such difficulty is that the symmetric-extension technique, which has been used to preserve sample rate in the separable subband system



of Fig. 1, does not extend to nonseparable filter systems. Another difficulty is that although a one-level nonseparable subband decomposition could have good spectral and directional sensitivity, further levels of subband decomposition would not necessarily produce subbands with useful spectral or directional properties. In addition to these difficulties, the computational complexity associated with nonseparable filtering would put any such approach at a disadvantage to our proposed approach, for which, as already noted, the arithmetic complexity of the receiver need not be substantially higher than that for a separable subband receiver.

It is our view that the orientation-adaptive algorithm presented in this paper provides highly competitive image qualities in the low bit rate arena. It is also open to the incorporation of more elaborate coding schemes that exploit dependencies between different subbands. Application to higher dimensional environments such as video remains an area for investigation.

## REFERENCES

- [1] E. H. Adelson, E. Simoncelli, and R. Hingorani, "Orthogonal pyramid transforms for image coding," in *Proc. SPIE*. (Cambridge, MA), Oct. 1987, pp. 50-58, vol. 845.
- [2] R. H. Bamberg and M. Smith, "A filter bank for the directional decomposition of images: Theory and design," *IEEE Trans. Signal Processing*, vol. 40, no. 4, pp. 882-893, Apr. 1992.
- [3] ———, "A comparison of directionally-based and nondirectionally-based subband image coders," in *Proc. SPIE, VCIP Conf. No. 1605* (Boston), Nov. 1991, pp. 757-768, vol. 2.
- [4] E. Chang and A. Zakhor, "Subband video coding based on velocity filters," in *Proc. IEEE Int. Symp. Circuits Syst.* (San Diego), 1992, pp. 2288-2291, vol. 5.
- [5] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 713-718.
- [6] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Amer. A*, vol. 2, no. 7, pp. 1160-1169, July 1985.
- [7] ———, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 36, no. 7, pp. 1169-1179, July 1988.
- [8] P. Delsarte, B. Macq, and D. T. M. Sloock, "Efficient multiresolution signal coding via a signal-adapted perfect reconstruction filter pyramid," in *Proc. Int. Conf. Acoust. Speech Signal Processing* (Toronto), 1991, pp. 2633-2636.
- [9] F. O'Gorman and M. B. Clowes, "Finding picture edges through colinearity of feature points," *IEEE Trans. Comput.*, vol. C-25, no. 4, pp. 449-456, Apr. 1976.
- [10] G. C. Gurski, M. T. Orchard, and A. W. Hull, "Optimal linear filters for pyramidal decomposition," in *Proc. Int. Conf. Acoust. Speech Signal Processing* (San Francisco), 1992, pp. 633-636, vol. 4.
- [11] A. Ikononopoulos and M. Kunt, "High compression image coding via directional filtering," *Signal Processing*, vol. 8, no. 2, pp. 179-203, April 1985.
- [12] H. Li and Z. He, "Directional subband coding of images," *International Conference on Acoustics, Speech and Signal Processing*, Glasgow, U.K., vol. 3, pp. 1823-1826, May 1989.
- [13] B. Mahesh and W. A. Pearlman, "Hexagonal sub-band coding for images," *International Conference on Acoustics, Speech and Signal Processing*, Glasgow, U.K., vol. 3, pp. 1953-1956, May 1989.
- [14] D. Marr and E. Hildreth, "Theory of Edge Detection," *Proc. R. Soc. Lond.* vol. B 207, pp. 187-217, 1980.
- [15] K. Nayeibi, T. Barnwell and M. Smith, "Analysis-synthesis systems with time-varying filter bank structures," *Proc. Int. Conf. Acoust. Speech Signal Processing* (San Francisco) 1992, pp. 617-620, vol. 4.
- [16] R. P. Rao and W. A. Pearlman, "Pyramid coding of images using visual criterion," in *Proc. SPIE (Visual Commun. Image Procession IV)* (Philadelphia), Nov. 1989, pp. 1269-1278, vol. 1199, pt. 3.
- [17] R. Rinaldo, D. Taubman, and A. Zakhor, "Applications of multiresolution analysis to images," in *Proc. Seventh Workshop Multidimensional Signal Processing* (Lake Placid, NY), Sept. 1991.
- [18] J. M. Shapiro, "Adaptive multidimensional perfect reconstruction filter banks using McClellan transformations," in *Proc. IEEE Int. Symp. Circuits Syst.* (San Diego) 1992, pp. 939-942, vol. 2.
- [19] F. P. Simoncelli and E. H. Adelson, "Subband Transforms," in *Subband Image Coding* (J. Woods, Ed.). Boston: Kluwer, 1990, pp. 143-192.
- [20] ———, "Nonseparable extensions of quadrature mirror filters to multiple dimensions," *Proc. IEEE*, vol. 78, no. 4, pp. 652-664, Apr. 1990.
- [21] M. J. T. Smith and S. L. Eddins, "Analysis-synthesis techniques for subband image coding," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 38, pp. 1446-1456, Aug. 1990.
- [22] D. Taubman and A. Zakhor, "A multistart algorithm for signal adaptive subband systems," in *Proc. Int. Conf. Acoust. Speech Signal Processing* (San Francisco), 1992, pp. 213-216, vol. 3.
- [23] P. P. Vaidyanathan and Z. Dogata, "The role of lossless systems in modern digital signal processing: A tutorial," Special Issue on Circuits and Systems, *IEEE Trans. Educat.*, Aug. 1989.
- [24] A. B. Watson, "The cortex transform: Rapid computation of simulated neural images," *Comput. Vision Graphics Image Processing*, vol. 39, pp. 311-327, 1987.



**David Taubman** received the B.S. degree in computer science and mathematics in 1986 and the B.E. degree in electrical engineering in 1988, both from the University of Sydney, Australia. He received the M.S. degree in electrical engineering in 1992 from the University of California at Berkeley and is currently enrolled in the Ph.D. program there.

From 1988 to 1990, he worked for the Electricity Commission of N.S.W., Australia. His research interests include subband image and video compression technology.

Mr. Taubman was awarded the University Medal from the University of Sydney in 1988 for electrical engineering. In the same year, he also received the Institute of Engineers, Australia, Prize and the Texas Instruments Prize for Digital Signal Processing, Australasia.



**Avidah Zakhor** (M'87) received the B.S. degree from the California Institute of Technology, Pasadena, and the S.M. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, all in electrical engineering, in 1983, 1985, and 1987, respectively.

In 1988, she joined the faculty at U. C. Berkeley, where she is currently Assistant Professor in the Department of Electrical Engineering and Computer Sciences. Her research interests are in the general area of signal processing and its applications to images and video as well as biomedical data. She has been a consultant to a number of industrial organizations and holds four U.S. patents.

Ms. Zakhor was a General Motors scholar from 1982 to 1983, received the Henry Ford Engineering Award and Caltech Prize in 1983, was a Hertz Fellow from 1984-1988, received the Presidential Young Investigator (PYI) Award, an IBM junior faculty development award, and an Analog Devices junior faculty development award in 1990, and the Office of Naval Research (ONR) Young Investigator Award in 1992. She is currently an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING and a member of the technical committee for multidimensional digital signal processing.