Non-Separable Extensions of Quadrature Mirror Filters to Multiple Dimensions

EERO P. SIMONCELLI, STUDENT MEMBER, IEEE AND EDWARD H. ADELSON, MEMBER, IEEE

Quadrature Mirror Filter (QMF) banks have been used in a variety of one-dimensional signal processing applications, and have been applied separably in two dimensions. As with most one-dimensional filters, separable extension to multiple dimensions produces a transform in which the orientation selectivity of some of the high-pass filters is poor. We describe generalized non-separable extensions of QMF banks to two and three dimensions, in which the orientation specificity of the high-pass filters is greatly improved. In particular, we discuss extensions to two dimensions with hexagonal symmetry, and three dimensional spatio-temporal extensions with rhombic-dodecahedral symmetry. Although these filters are conceived and designed on non-standard sampling lattices, they may be applied to rectangularly sampled images. As in one dimension, these transformations may be hierarchically cascaded to form a multi-scale "pyramid" representation. We design a set of example filters and apply them to the problems of image compression, progressive transmission, orientation analysis, and motion analysis.

INTRODUCTION

Sub-band transforms have been successfully employed in many areas of signal processing. For many applications, especially in image processing, researchers have advocated the use of sub-band transforms that divide the frequency spectrum into octave bandwidth pieces [1]-[4]. In such a transform the basis functions represent information at spatial scales which are related by powers of two.

A particularly useful one-dimensional orthogonal sub-band transform is the Quadrature Mirror Filter (QMF) bank, which was introduced by Croisier et al. [5], [6]. These filters are used in an analysis/synthesis system which decomposes a signal into high-pass and low-pass frequency sub-bands. They are also well-suited for octave band splitting, since they can be applied recursively to split the low-pass sub-band. Vaidyanathan developed a more general theory of

perfect reconstruction filter banks based on a polyphase matrix decomposition in the frequency domain [7]. Vetterli was the first to propose the use of QMFs for image decomposition [8]. He showed examples of both separable and non-separable non-oriented QMF decompositions in two dimensions. Vaidyanathan established criteria for perfect reconstruction QMF banks for two-dimensional applications [9]. Viscito and Allebach developed perfect reconstruction multi-dimensional filter banks with arbitrary decimation patterns [10]. Woods and O'Neil used separable QMFs for image data compression [11]. Several other authors have used QMF pyramid transforms for data compression [12]-[14]. Mallat [3] related QMF pyramids to wavelet theory and proposed their use in machine vision.

In a parallel development, there has been a great deal of interest in image representations that are tuned for orientation as well as scale. This is equivalent to requiring that the frequency spectra of the basis functions of the representation exhibit angular localization. A variety of arguments have been advanced in favor of such transforms, based on properties of the human visual system and the statistics of images [15]-[18]. Daugman, and Porat et al. have explored two-dimensional Gabor transforms, in which the basis functions are Gaussian windows modulated by sinusoidal gratings [18], [19]. A related transform has been described by Watson [16].

Most applications of QMFs to two or more dimensions have involved separable filters. A two-dimensional example is illustrated in Fig. 1: The frequency spectrum is split into low-pass, horizontal high-pass, vertical high-pass, and diagonal high-pass sub-bands. The diagonal band contains mixed orientations.

Adelson et al. [12] demonstrated that one could develop a non-separable decomposition based on hexagonally symmetric QMFs, thereby achieving an orthogonal transform in which the all of the basis functions are localized in space, spatial frequency, and orientation. This approach to pyramid construction allows smooth spatial overlap between basis functions, and produces much better frequency tuning than does the blocked non-overlapping construction employed by Crettez and Simon [20] or Watson and Ahumada [21]. In the present paper we extend the Adelson et al. concepts; we describe methods for designing hexagonal QMFs, and we apply them to a variety of problems. We also
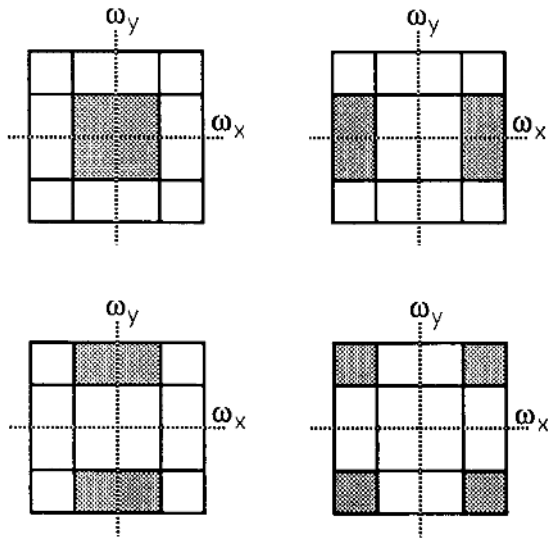
Fig. 1. Idealized partition of the frequency domain by separable application of two-band one-dimensional QMFs.

describe three-dimensional generalizations with rhombic-dodecahedral symmetry.

REVIEW OF ONE DIMENSIONAL QMF CONCEPTS

In this section, we give a brief review of Quadrature Mirror Filters in one dimension. A more thorough review may be found in [22], or more recently, [23] or [24]. The original QMF problem was formulated as a two-band critically sampled analysis/synthesis filter bank problem, as illustrated in the schematic diagram in Fig. 2. The purpose of the analysis section of the filter bank is to decompose the input sequence $x[n]$ into two half-density representation sequences $y_0[n]$ and $y_1[n]$. The synthesis section then recombines these sequences to form an approximation $\hat{x}[n]$ to the original sequence. The system is called "critically sampled" because the sample input rate is equal to the total sample rate of the intermediate sequences. The notation in the diagram is standard for digital signal processing. The boxes $\boxed{F_i(\omega)}$ indicate convolution of an input sequence with a filter with impulse response $f_i[n]$ and discrete time Fourier transform (DTFT)
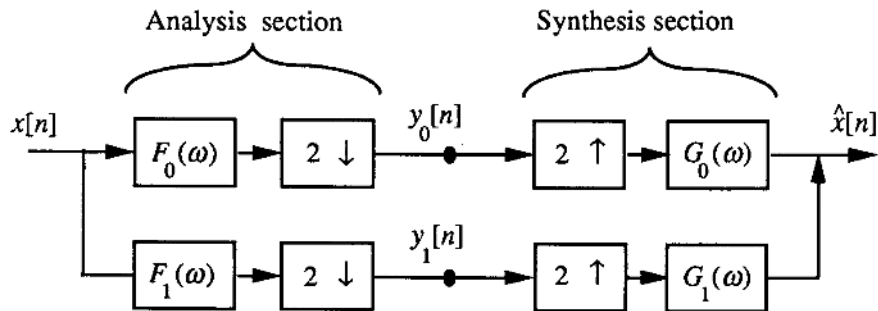
$$F_i(\omega) = \sum_n f_i[n]e^{-j\omega n}$$

The boxes $\boxed{2\downarrow}$ indicate that the sequence is subsampled by a factor of 2, and the boxes $\boxed{2\uparrow}$ indicate that the sequence should be upsampled by inserting a zero between each sample.

Using the definition of the DTFT and some well known facts about the effects of upsampling and downsampling in the frequency domain, one can derive equations for the DTFT of the representation sequences $y_i[n]$:

$$Y_i(\omega) = \frac{1}{2}\left[F_i\left(\frac{\omega}{2}\right)X\left(\frac{\omega}{2}\right) + F_i\left(\frac{\omega}{2} + \pi\right)X\left(\frac{\omega}{2} + \pi\right)\right] \quad (1)$$

and the A/S system output is

$$\hat{X}(\omega) = Y_0(2\omega)G_0(\omega) + Y_1(2\omega)G_1(\omega).$$

Combining these equations gives the overall system response of the filter bank:

$$\hat{X}(\omega) = \frac{1}{2}[F_0(\omega)G_0(\omega) + F_1(\omega)G_1(\omega)]X(\omega)$$

$$+ \frac{1}{2}[F_0(\omega + \pi)G_0(\omega) + F_1(\omega + \pi)G_1(\omega)]X(\omega + \pi). \quad (2)$$

The first term is a linear shift-invariant (LSI) system response, and the second is the system aliasing.

The term QMF refers to a clever choice of filters that are related by spatial shifting and frequency modulation. We define

$$F_0(\omega) = G_0(-\omega) = H(\omega)$$

$$F_1(\omega) = G_1(-\omega) = e^{j\omega}H(-\omega + \pi) \quad (3)$$

for $H(\omega)$ an arbitrary function of ω. This definition, which was proposed in [25], corresponds to the linear algebraic notion of an orthogonal transform, and is a more general definition than that originally provided by Croisier et al. In particular, the original definition does not contain an explicit spatial (temporal) shift factor and is therefore valid only for even-length filters. The definition given above contains the original as a subcase, and is also valid for odd-length filters.

With the choice of filters given in (3), equation (2) becomes

$$\hat{X}(\omega) = \frac{1}{2}[H(\omega)H(-\omega) + H(-\omega + \pi)H(\omega + \pi)]X(\omega)$$

$$+ \frac{1}{2}[H(\omega + \pi)H(-\omega)$$

$$+ e^{j\pi}H(-\omega)H(\omega + \pi)]X(\omega + \pi). \quad (4)$$

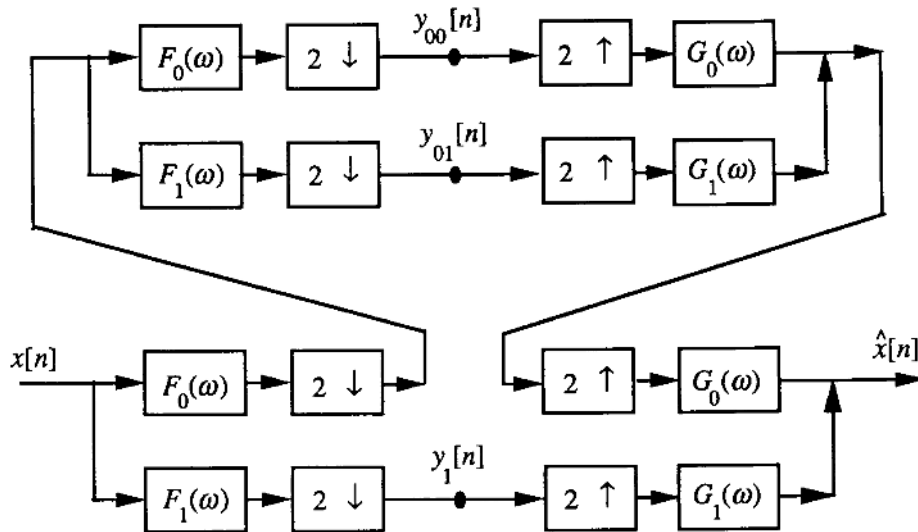**Fig. 2.** A two-band analysis/synthesis filter bank in one dimension.

Fig. 3. A non-uniformly cascaded analysis/synthesis filter bank.

The second (aliasing) term cancels, and the remaining LSI system response is

$$\hat{X}(\omega) = \tfrac{1}{2}[H(\omega)\, H(-\omega) + H(-\omega + \pi)\, H(\omega + \pi)]\, X(\omega). \quad (5)$$

Note that the aliasing cancellation is exact, independent of the choice of the function $H(\omega)$. We should emphasize, however, that it is the overall system aliasing that cancels— the individual sub-bands *do* contain aliasing.

The design problem is now reduced to finding a filter with DTFT $H(\omega)$ that satisfies the constraint

$$\tfrac{1}{2}[H(\omega)H(-\omega) + H(-\omega + \pi)H(\omega + \pi)] = 1$$

or

$$|H(\omega)|^2 + |H(\omega + \pi)|^2 = 2. \quad (6)$$

In general, lowpass solutions are desirable since the system then becomes a band-splitting system. Several authors have studied the design and implementation of these filters [25]–[29]. Once filters have been designed so that the overall system response is unity, the filter bank may be cascaded to form multiple-band systems. This may be done in a uniform manner as in [11], or in a non-uniform or "pyramid" fashion [12]. An example of this non-uniform cascading is illustrated in Fig. 3. Such a pyramid cascade produces an octave-width sub-band decomposition, as illustrated in the idealized frequency diagram in Fig. 4.

TWO-DIMENSIONAL HEXAGONAL FILTERS

As discussed in the introduction, separable application of one-dimensional QMFs produces a representation in which one of the sub-bands contains a mixture of two orientations. The problem is inherent in the rectangular sampling scheme: The frequency response of any rectangularly sampled function has the same value at the points (π, π), $(-\pi, \pi)$, $(\pi, -\pi)$, and $(-\pi, -\pi)$, so this frequency does not correspond to a pure orientation. Splitting the frequencies in the neighborhood of this point into different orientation
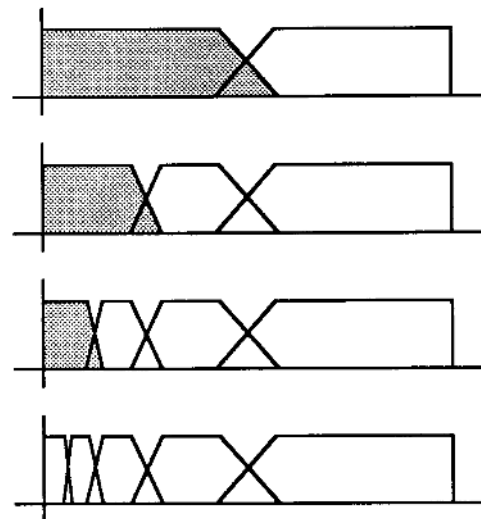

Fig. 4. Octave band splitting produced by a four-level pyramid cascade of a two-band A/S system. The top picture represents the splitting of the two-band A/S system. Each successive picture shows the effect of reapplying the system to the lowpass sequence (indicate in grey) of the previous picture. The bottom picture gives the final four-level partition of the frequency domain. All frequency axes cover the range from 0 to π.

bands requires the use of very large filters. In general, the high-frequency diagonal regions of the spectra of real images are relatively insignificant. But if the filter bank is cascaded to form a pyramid, then the lower frequency diagonals (where there *is* significant power) will also be mixed.

In this section, we will discuss the use of hexagonal sampling systems and filters. Hexagonal sampling lattices provide the tightest packing of all regular two-dimensional lattices. In addition, we will show that the mixed orientation problem discussed above can be eliminated by using hexagonally symmetric filters.
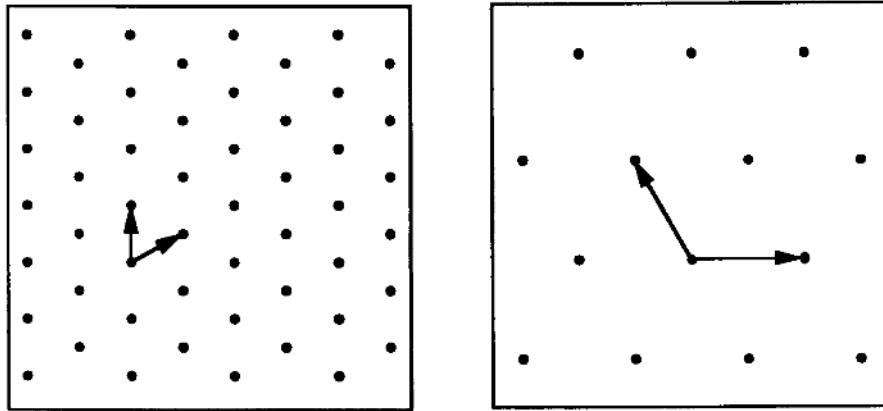
Fig. 5. Relationship between hex sampling lattices in the spatial and spatial frequency domains. On the left is the lattice defined by the sampling vectors. On the right is the Fourier transform of this lattice, defined by the modulation vectors.

Figure 5 shows a hexagonal sampling lattice and its Fourier transform. The sampling lattice is defined by a pair of *sampling vectors* in the plane:

$$v_0 = \begin{pmatrix} \sqrt{3}/2 \\ 1/2 \end{pmatrix}, \qquad v_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The locations of the lattice points consist of all linear combinations of these vectors with integer coefficients. In the frequency domain, the effect of this sampling is to convolve the original frequency spectrum of the image with a *modulation* or *reciprocal lattice* which is the Fourier transform of the sampling lattice. The modulation lattice is defined by a pair of *modulation vectors* in the frequency plane:

$$\tilde{v}_0 = \begin{pmatrix} 4\pi/\sqrt{3} \\ 0 \end{pmatrix}, \qquad \tilde{v}_1 = \begin{pmatrix} -2\pi/\sqrt{3} \\ 2\pi \end{pmatrix}.$$

Thus if $H(\omega)$ is the Fourier transform of a hexagonally sampled signal (image) then it is invariant to translations by multiples of the vectors $\tilde{v}_i$:

$$H(\omega) = H(\omega + n_0\tilde{v}_0 + n_1\tilde{v}_1) \qquad (7)$$

for n_0 and n_1 any two integers.

In general, the relationship between the sampling vectors and modulation vectors is easily described in terms of matrices [30]–[31]. If we consider the *sampling matrix* V with columns containing the vectors v_i and the *modulation matrix* $\tilde{V}$ with columns containing the vectors $\tilde{v}_i$, then the two matrices are related by the equation

$$\tilde{V} = 2\pi(V^{-1})^t. \qquad (8)$$

Note that we know V is invertible since we assume that the sampling vectors span the space and are therefore linearly independent.

Next, we need to consider the analysis/synthesis filter bank in multiple dimensions. A four-band A/S filter bank in two dimensions is illustrated in Fig. 6. It is similar to its one-dimensional counterpart in Fig. 2, but the filtering and subsampling is done in two dimensions: ω is now a two-dimensional vector, and the subsampling is parameterized by a non-singular two-by-two *subsampling matrix*, K, with integer entries.

Figure 7 illustrates two-dimensional subsampling in both the spatial and frequency domains. We can think of the subsampling operation as sampling the original continuous
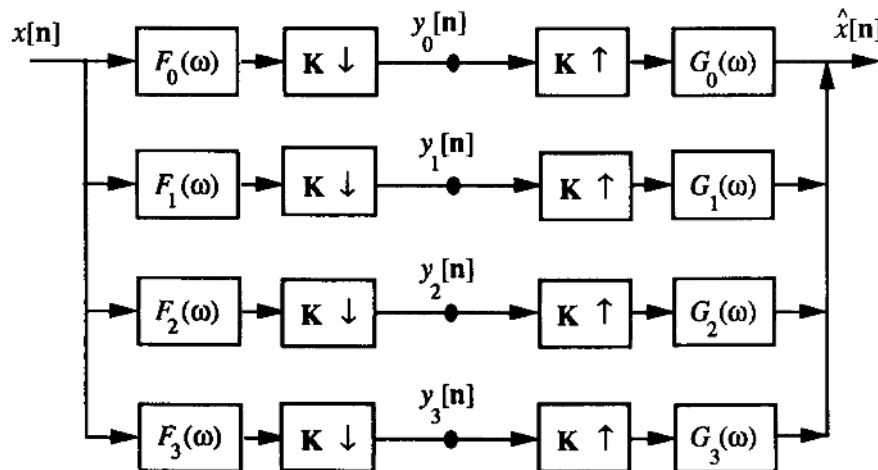

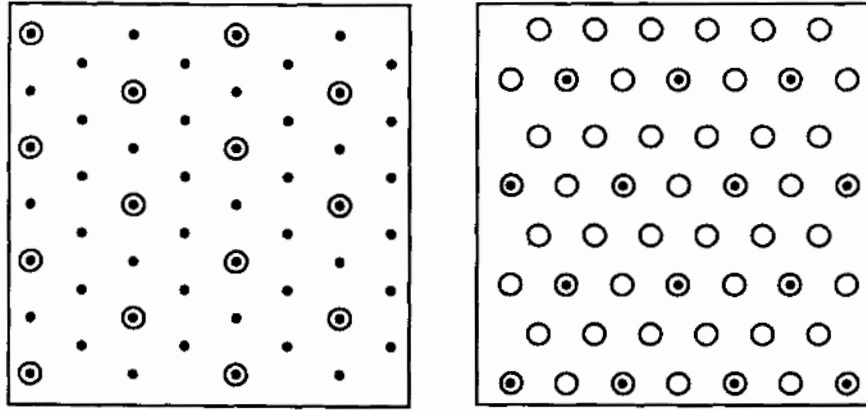**Fig. 6.** A two-dimensional four-band analysis/synthesis filter bank.

Fig. 7. Illustration of subsampling on a hexagonal lattice. The points in the diagram on the left represent the original sampling lattice and the circles represent the subsampled lattice points. The picture on the right shows the Fourier transform of the lattice (points) and the Fourier transform of the subsampled lattice (circles).

image on a lattice defined by two vectors u_i which are linear combinations of the sampling vectors v_j. Then the new sampling matrix is written as

$$U = VK.$$

In order to write a general expression for the output of a multi-dimensional analysis/synthesis system, we need an equation analogous to that given in (1) relating the subsampled signal to the sample signal in the frequency domain. For rectangular sampling lattices in d dimensions, the relationship is simple. The sampling matrix K generates a sublattice defined by

$$\{n : n = Km, m \in Z^d\},$$

where Z^d is the set of all d-dimensional vectors with integer components. The sublattice has $|K|$ distinct cosets, each coset being a copy of the sublattice translated by an integer vector, and the union of the cosets is the original sampling lattice [10]. Consider two signals related by subsampling: $s[n] = r[Kn]$. Then their Fourier transforms are related by the expression

$$S(\omega) = \frac{1}{|K|} \sum_{i=0}^{|K|-1} R((K^{-1})^t(\omega - 2\pi k_i))$$

where $S(\omega)$ and $R(\omega)$ are the Fourier transforms of $s[n]$ and $r[n]$ respectively, and the k_i are a set of polyphase shift vectors corresponding to each of the $|K|$ sublattice cosets [10]. A simple example of a set of shift vectors is the following:

$$\{k : (K^{-1})^t k \in [0, 1)^d, k \in Z^d\}.$$

The corresponding expression for non-rectangular sampling lattices is obtained by mapping from the rectangular case. The result of subsampling in the analysis/synthesis system may then be written as a convolution of the sampled spectrum with a set of *subsampling modulation vectors* $\tilde{\kappa}_j$:

$$Y_i(\omega) = \frac{1}{|K|} \sum_{j=0}^{|K|-1} F_i((K^{-1})^t\omega - \tilde{\kappa}_j)X((K^{-1})^t\omega + \tilde{\kappa}_j), \quad (9)$$

where the $\tilde{\kappa}_j$ are defined as

$$\{\tilde{\kappa}_j : j = 0, 1, \cdots, |K| - 1\}$$
$$= \{\tilde{V}(K^{-1})^t n : (K^{-1})^t n \in [0, 1)^d, n \in Z^d\}. \quad (10)$$

The effect of upsampling in the frequency domain is the same as for the rectangular case [31]. Combining equation (9) with the frequency domain upsampling relationship gives an expression for the overall filter bank response:

$$\hat{X}(\omega) = \frac{1}{|K|} \sum_{i=0}^{|K|-1} G_i(\omega) Y_i(K^t\omega)$$

$$= \frac{1}{|K|} \sum_{i=0}^{|K|-1} G_i(\omega) \sum_{j=0}^{|K|-1} F_i(\omega + \tilde{\kappa}_j)X(\omega + \tilde{\kappa}_j)$$

$$= \frac{1}{|K|} \sum_{j=0}^{|K|-1} X(\omega + \tilde{\kappa}_j) \left[\sum_{i=0}^{|K|-1} G_i(\omega)F_i(\omega + \tilde{\kappa}_j) \right]. \quad (11)$$

As in equation (2), one term of the sum corresponds to the LSI system response, and the remaining terms are the system aliasing.

Returning now to the specific case of the hexagonal sampling lattice, we wish to select a sampling matrix K. Since we want to be able to apply our transform recursively, we choose a subsampling scheme which preserves the geometry of the original sampling lattice:

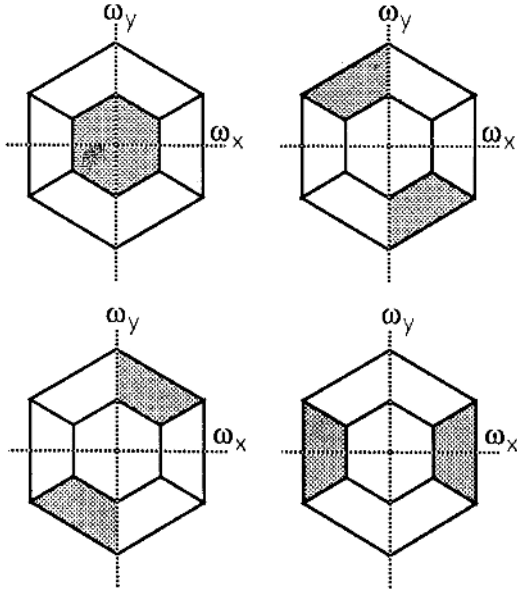$$K = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

On the hexagonal sampling lattice with this subsampling scheme, the definition given in (10) produces the following modulation vectors:

$$\tilde{\kappa}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \tilde{\kappa}_1 = \begin{pmatrix} 2\pi/\sqrt{3} \\ 0 \end{pmatrix},$$

$$\tilde{\kappa}_2 = \begin{pmatrix} \pi/\sqrt{3} \\ \pi \end{pmatrix}, \qquad \tilde{\kappa}_3 = \begin{pmatrix} -\pi/\sqrt{3} \\ \pi \end{pmatrix}.$$

Figure 8 offers an idealized picture of this modulation.

Now we are ready to define a set of hexagonal QMFs. Analogous to the one-dimensional case, we can choose the filters to eliminate the aliasing terms in equation (11):

$$F_0(\omega) = G_0(-\omega) = H(\omega) = H(-\omega)$$
$$F_1(\omega) = G_1(-\omega) = e^{j\omega \cdot s_1} H(\omega + \tilde{\kappa}_1)$$
$$F_2(\omega) = G_2(-\omega) = e^{j\omega \cdot s_2} H(\omega + \tilde{\kappa}_2)$$
$$F_3(\omega) = G_3(-\omega) = e^{j\omega \cdot s_3} H(\omega + \tilde{\kappa}_3) \qquad (12)$$

Fig. 8. Illustration of the modulating effect of subsampling in the frequency domain. Assume that the sampled image has a spectrum bandlimited to the gray region in the upper left frequency diagram. Subsampling will modulate the spectrum to the gray regions in the other three diagrams. The resulting spectrum will be the sum of the four spectra.

where H is a function that is invariant under negation of its argument, and the expressions $\omega \cdot s_i$ indicates an inner product of the two vectors. As in equation (3), the filters are related by spatial shifting and frequency modulation. For the subsampling matrix we are using here, there are four sublattice cosets and therefore four distinct shifting vectors (including the zero vector). Two assignments of the s_i lead to system aliasing cancellation, and these two assignments are related by reflection through the origin. So without loss of generality, we choose the shifting vectors to be

$$s_1 = \begin{pmatrix} \sqrt{3}/2 \\ 1/2 \end{pmatrix}, \quad s_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad s_3 = \begin{pmatrix} -\sqrt{3}/2 \\ 1/2 \end{pmatrix}$$

To see that the choice of filters in (12) forces cancellation of the aliasing terms, consider the aliasing response (the bracketed expression) in the second term ($j = 1$) of equation (11):

$$\sum_{i=0}^{3} G_i(\omega) F_i(\omega + \bar{\kappa}_1)$$

$$= H(-\omega)H(\omega + \bar{\kappa}_1)$$

$$+ e^{-j\omega \cdot s_1} H(\omega + \bar{\kappa}_1) \cdot e^{j(\omega + \bar{\kappa}_1) \cdot s_1} H(\omega + 2\bar{\kappa}_1)$$

$$+ e^{-j\omega \cdot s_2} H(-\omega + \bar{\kappa}_2) \cdot e^{j(\omega + \bar{\kappa}_1) \cdot s_2} H(\omega + \bar{\kappa}_1 + \bar{\kappa}_2)$$

$$+ e^{-j\omega \cdot s_3} H(-\omega + \bar{\kappa}_3) \cdot e^{j(\omega + \bar{\kappa}_1) \cdot s_3} H(\omega + \bar{\kappa}_1 + \bar{\kappa}_3)$$

$$= H(-\omega)H(\omega + \bar{\kappa}_1)$$

$$+ e^{j\bar{\kappa}_1 \cdot s_1} H(-\omega + \bar{\kappa}_1)H(\omega)$$

$$+ e^{j\bar{\kappa}_1 \cdot s_2} H(-\omega + \bar{\kappa}_2)H(\omega + \kappa_3)$$

$$+ e^{j\bar{\kappa}_1 \cdot s_3} H(-\omega + \bar{\kappa}_3)H(\omega + \bar{\kappa}_2)$$

$$= 0.$$

We have made use of the assumed symmetry of $H(\omega)$ and the translation invariance defined in equation (7). The terms for $j = 2$ and $j = 3$ cancel in a similar manner.

After cancelling all of the aliasing terms in equation (11), the remaining LSI system response is

$$\hat{X}(\omega) = \frac{1}{4} X(\omega) \sum_{i=0}^{3} G_i(\omega) F_i(\omega)$$

$$= \frac{1}{4} X(\omega) \sum_{i=0}^{3} H(-\omega + \bar{\kappa}_i) H(\omega + \bar{\kappa}_i)$$

$$= \frac{1}{4} X(\omega) \sum_{i=0}^{3} |H(\omega + \bar{\kappa}_i)|^2. \tag{13}$$

As in one dimension, the aliasing cancellation is exact, independent of the choice of $H(\omega)$, and the design problem is reduced to finding a filter with DTFT $H(\omega)$ satisfying the constraint

$$\sum_{i=0}^{3} |H(\omega + \bar{\kappa}_i)|^2 = 4. \tag{14}$$

This is analogous to the one-dimensional equation (6). Again, a low-pass solution will produce a band-splitting system which may be cascaded hierarchically to produce an octave bandwidth decomposition in two dimensions. An idealized illustration of this is given in Fig. 9. Figure 10 shows the results of applying a hexagonal QMF bank to an image of a disk.

Hexagonal Filter Design

A "good" filter is one that satisfies the constraint given in equation (14). In addition, many applications require that the sub-band images have a minimal amount of aliasing (as in the one-dimensional case, the individual subbands contain aliasing, although the overall system aliasing cancels). Our objective was to design filters with small regions of support that satisfy both of these constraints. In one dimension, we found that a simple frequency-sampling design method produced high quality QMFs with small regions of support [25]. We have found that this method is suitable for multi-dimensional designs as well.

We begin by fixing the size of the hexagonal filter kernel and the sampling density in the frequency domain. The size of the kernel is measured in terms of the number of hexagonal "rings" it contains. For example, a zero-ring filter contains only a single impulse, and a one-ring filter contains a center impulse surrounded by a hexagonal ring of six more impulses. Due to the hexagonal symmetry, the free parameters of the filter comprise a wedge-shaped region covering approximately one twelfth of the kernel. These parameters completely specify the kernel.

We define a filter bank error function as the maximal deviation of the overall filter bank response given in equation (13) from its ideal value:

$$E_1 = \max_{\omega} \left\{ f_1(\omega) \left| \sum_{i=0}^{3} |H(\omega + \bar{\kappa}_i)|^2 - 4 \right| \right\}$$

where ω ranges over the samples in the frequency spectrum. The function $f_1(\omega)$ is a frequency weighting function roughly matched to the sensitivity of the human visual system and the statistics of images:
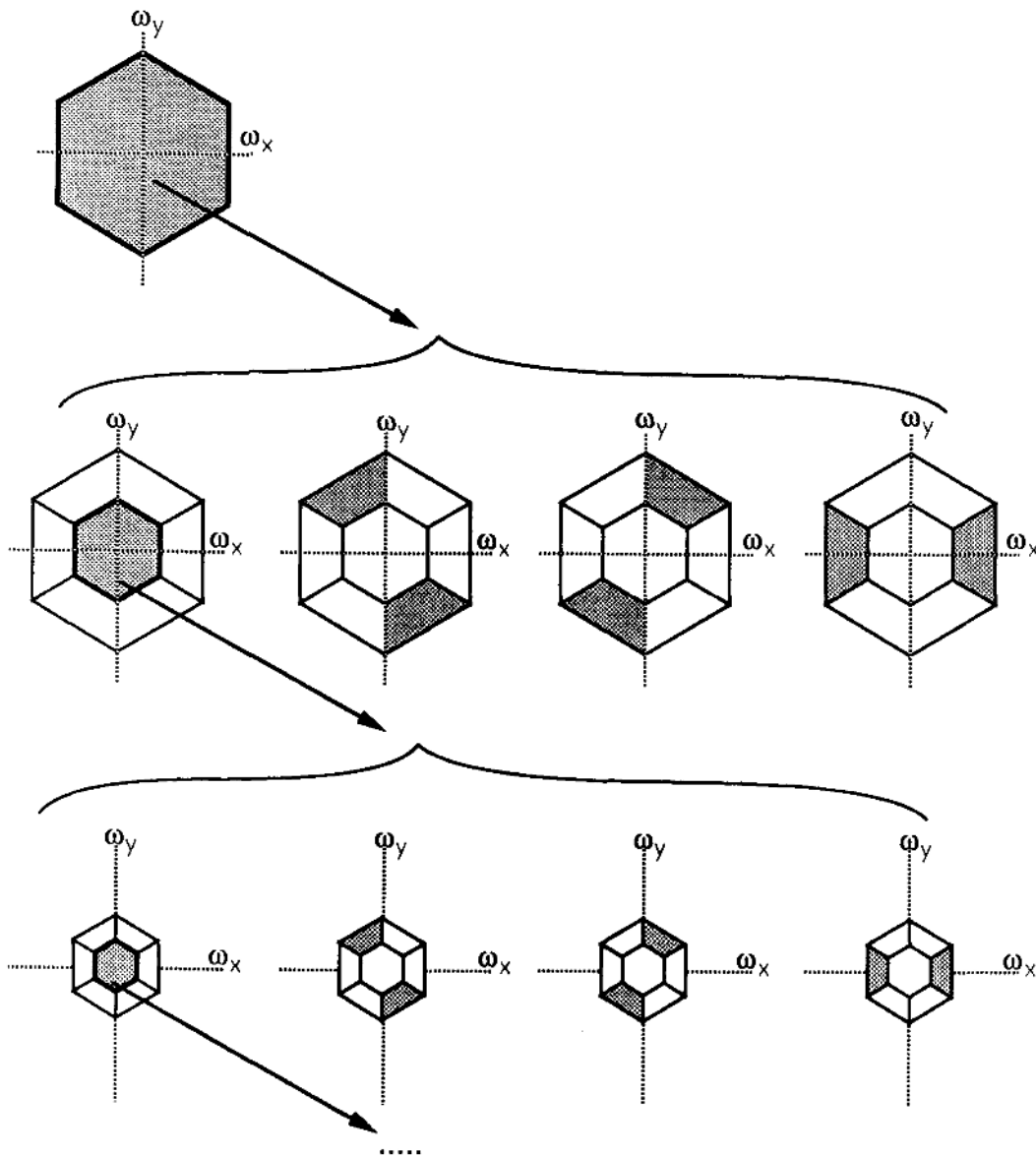
$$f_1(\omega) = 1/|\omega|.$$

Fig. 9. Idealized diagram of the partition of the frequency plane resulting from a four-level pyramid cascade of hexagonal filters. The top plot represents the frequency spectrum of the original image. This is divided into four sub-bands at the next level. On each subsequent level, the lowpass sub-band (outlined in bold) is sub-divided further.

We also define an intra-band aliasing error function:

$$E_2 = \max_{\omega'} \{ f_2(\omega') | H(-\omega')H(\omega' + \check{\kappa}_1)| \}$$

where $\check{\kappa}_i$ is any one of the non-zero subsampling modulation points and the function $f_2(\omega')$ is defined as

$$f_2(\omega') = 1/|\omega'|^2.$$

The frequency vector ω' ranges over all of the samples in the frequency spectrum, except for those in a hexagonal boundary containing the point $(0, \pi/2)$. Aliasing within sub-bands cannot be eliminated at the points in this boundary because the overall filter bank response at these points

would then be forced to zero, violating the constraint in equation (14).

Finally, we combine the two error functions as a weighted sum:

$$E = \alpha E_1 + (1 - \alpha)E_2, \quad \alpha \in [0, 1].$$

Given a set of values for the free parameters, we can construct a kernel and compute the value of the error E. To design filters, we used a downhill simplex method to search the space of free parameters for minima in E. The weighting factor α was adjusted to give a filter bank response error E_1 less than a fixed threshold. The power spectra of an exam-
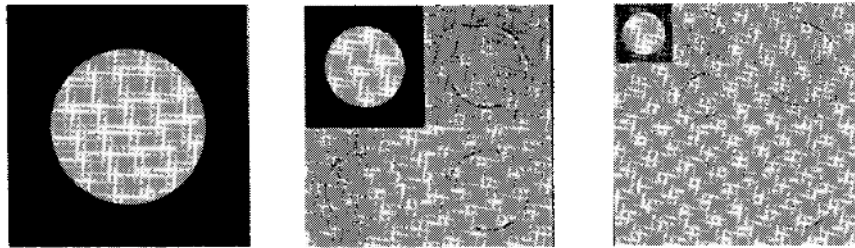
Fig. 10. Results of applying a hexagonal QMF bank to an image of a disk. On the left is the original image. In the center is the result after one application of the analysis section of the filter bank. The image has been decomposed into a low-pass and three oriented high-pass images at 1/4 density. On the right, we have applied the filter bank recursively to the low-pass image to produce a two-level pyramid decomposition.

ple set of filters is plotted in Fig. 11. Several example filter kernels are given in the Appendix.
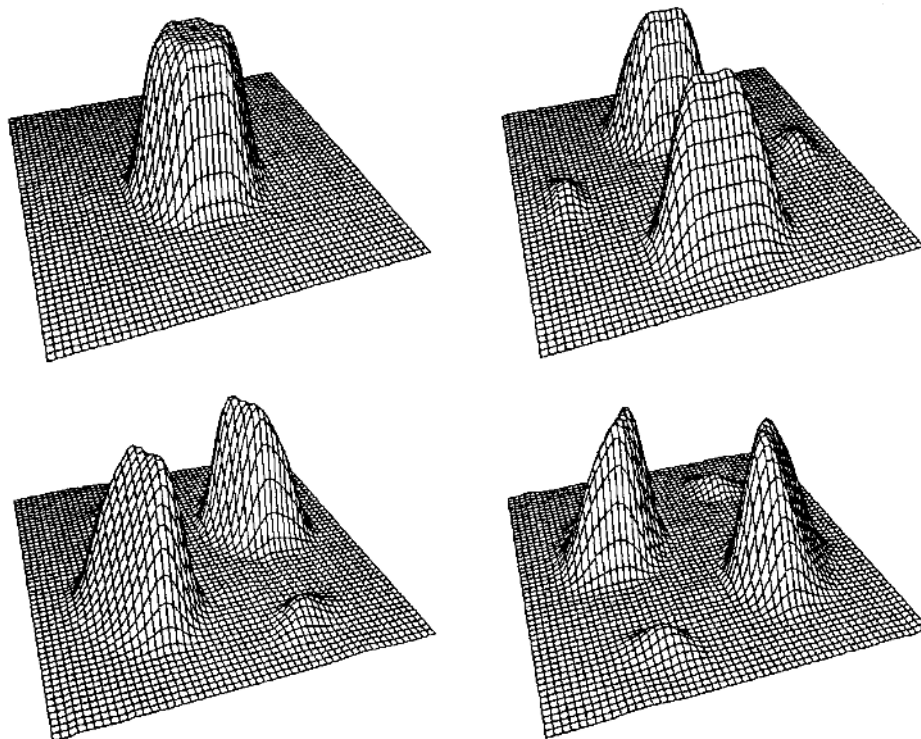
Image Data Compression

We compressed the standard "Lena" image using the "4-ring" hexagonal filter given in the Appendix. The original image is shown in Fig. 12 and two compressed versions in Fig. 13 and Fig. 14. To hexagonally sample the image, we resampled vertically by a factor of 7/4 using sinc interpolation. We then multiplied by the function $f[n] = 1 + (-1)^{(n_x + n_y)}$. This method, which is similar to one suggested by Mersereau [30], gives a reasonable geometric approximation to a hexagonal sampling lattice.

A four-level pyramid transform was applied to the image.

This transformation partitions the frequency domain into octave-spaced oriented sub-bands, as illustrated in the idealized frequency diagram of Fig. 9. An overall bit rate (entropy) R was fixed and the bit rates assigned to the coefficients of the transform were determined using the standard optimal allocation formula [32]:

$$R_k = R = \frac{1}{2} \log_2 \frac{\sigma_k^2}{\left[\prod_{j=0}^{N-1} \sigma_j^2 \right]^{1/N}}$$

where σ_k^2 is the variance of the kth coefficient in the transform. Negative values of R_k were set to zero and the other bit rates raised to maintain the correct overall bit rate R. Each sub-image was quantized with the bin size chosen to

Fig. 11. The power spectra for the "4-ring" set of hexagonal QMF filters. The filter kernels are given in the appendix.

SIMONCELLI AND ADELSON: NON-SEPARABLE EXTENSIONS OF QUADRATURE MIRROR FILTERS

Fig. 12. The original "Lena" image at 256 × 256 pixels.

Fig. 13. Data compression using the 4-ring hexagonal filter bank. The entropy of the quantized pyramid was 1.0 bits per pixel for a total of 65 536 bits.

Fig. 14. Data compression using the 4-ring hexagonal filter bank. The entropy of the quantized pyramid was 0.25 bits per pixel for a total of 16 384 bits.

a gradual manner. In the case of a sub-band pyramid, this is easily accomplished by sending the information in order from lowest to highest resolution. Figure 15 shows a sequence of intermediate progressive transmission images for a four-level pyramid transform using hexagonal filters.
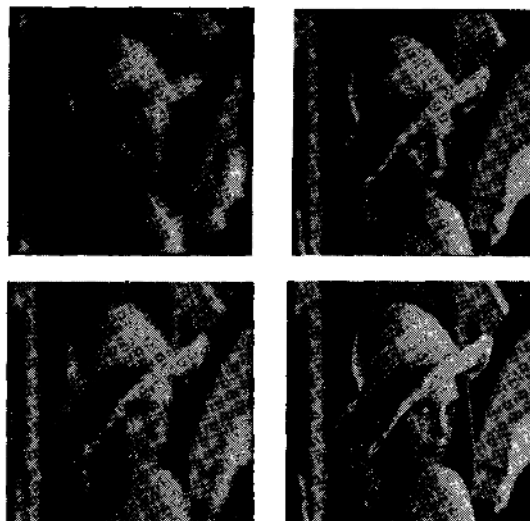
Fig. 15. Progressive transmission images using a four-level pyramid transform constructed with hexagonal filters. The first image is produced by reconstructing only the low-pass signal at the fourth pyramid level. The next image results from reconstructing the entire fourth level, or equivalently, the low-pass image on the third level. The final image is not shown.

give a first order entropy equal to the optimal bit rate R_k for that sub-image.

Upon re-synthesizing the image, we interpolated the zero-valued pixels and vertically resampled by a factor of 4/7. We feel that the compression results are superior to comparable results using a block DCT. The results also compare favorably with similar results using separable QMF pyramids: the aliasing errors are less disturbing visually than those of separable QMFs.

Progressive Transmission

Another useful application of pyramid sub-band representations is progressive transmission. This involves sending an image through a low-capacity channel in such a way that a low resolution version of the image becomes available quickly, and higher resolution information is added in

Orientation Analysis

To demonstrate the orientation selectivity of the filters, we use a technique developed by Freeman and Adelson [33]. Using a set of three second-order polynomial filters,

they extract the local orientation of the image at each pixel position. Although the hexagonal filters developed here are not second-order polynomials, they have similar frequency spectra and thus have approximately the same behavior when used in this fashion. Figure 17 shows a line drawing representation of the result of applying the technique to the image of Einstein given in Figure 16. The line orientation corresponds to the local image orientation, and the line length corresponds to the degree of local anisotropy: Longer lines indicate that the image is strongly oriented.

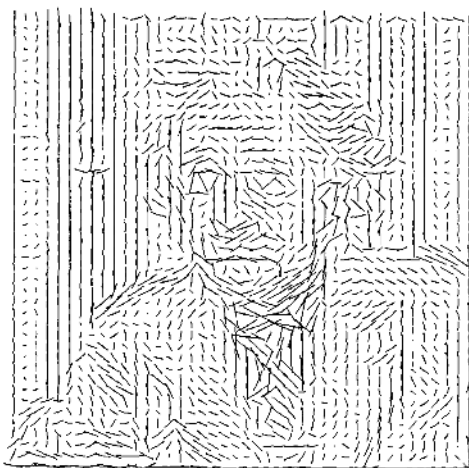Fig. 16. Original image of Einstein.

Fig. 17. An oriented line drawing of Albert Einstein, produced by using linear combinations of the oriented hexagonal filter outputs to measure the strength and orientation of the local image anisotropy.

RHOMBIC DODECAHEDRAL FILTERS IN THREE DIMENSIONS

The extension of the concepts developed in the previous section to three dimensions is fairly straightforward. Analogous to the two-dimensional hexagonal case, we choose a periodic sampling lattice which corresponds to the densest packing of spheres in three dimensions. This is the
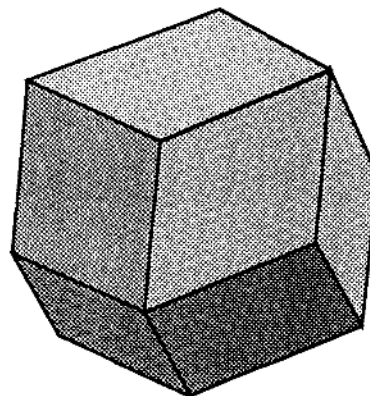

Fig. 18. A rhombic dodecahedron. This is the shape of the repeating frequency region for the "garnet" filter.

crystal structure of garnet, and therefore we have given these filters the nickname "garnet" QMFs. We choose as a band-limiting region the *Voronoi* region of this lattice (a rhombic dodecahedron) which is illustrated in Fig. 18. The sampling matrix for the lattice is

$$\mathbf{V} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix}.$$

Using equation (8), the modulation matrix is then

$$\tilde{\mathbf{V}} = \begin{bmatrix} \pi & 0 & 0 \\ -\pi & 2\pi & 0 \\ -\sqrt{2}\pi & 0 & 2\sqrt{2}\pi \end{bmatrix}.$$

To preserve the geometry of the original sampling lattice, we choose an eight-band A/S system with subsampling matrix

$$\mathbf{K} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

This produces the following subsampling modulation points, as determined by equation (10):

$$\tilde{\mathbf{\kappa}}_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \tilde{\mathbf{\kappa}}_1 = \begin{pmatrix} 0 \\ \pi \\ \sqrt{2}\pi \end{pmatrix},$$

$$\tilde{\mathbf{\kappa}}_2 = \begin{pmatrix} 0 \\ \pi \\ 0 \end{pmatrix}, \quad \tilde{\mathbf{\kappa}}_3 = \begin{pmatrix} \pi/2 \\ \pi/2 \\ -\pi/\sqrt{2} \end{pmatrix},$$

$$\tilde{\mathbf{\kappa}}_4 = \begin{pmatrix} \pi/2 \\ -\pi/2 \\ \pi/\sqrt{2} \end{pmatrix}, \quad \tilde{\mathbf{\kappa}}_5 = \begin{pmatrix} \pi/2 \\ -\pi/2 \\ -\pi\sqrt{2} \end{pmatrix},$$

$$\tilde{\mathbf{\kappa}}_6 = \begin{pmatrix} \pi/2 \\ \pi/2 \\ \pi/\sqrt{2} \end{pmatrix}, \quad \tilde{\mathbf{\kappa}}_7 = \begin{pmatrix} 0 \\ 0 \\ \sqrt{2}\pi \end{pmatrix}.$$

These modulation vectors correspond to a decomposition into the following sub-bands:

—low-pass
—stationary vertical
—stationary horizontal
—motion up/right
—motion up/left
—motion down/right
—motion down/left
—combined stationary diagonals and full-field flicker.

Unfortunately, there seems to be no way to avoid the last filter which contains mixed orientation. The overall system response of the filter bank is

$$\hat{X}(\omega) = \frac{1}{8} \sum_{j=0}^{7} X(\omega + \bar{\kappa}_j) \left[\sum_{i=0}^{7} G_i(\omega) F_i(\omega + \bar{\kappa}_j) \right]. \quad (15)$$

where the first term is the LSI system response, and the remaining terms are aliasing terms.

Once again, we can choose filters related by shifts and modulations that will cancel the system aliasing terms:

$$F_0(\omega) = G_0(-\omega) = H(\omega) = H(-\omega)$$

$$F_i(\omega) = G_i(-\omega) = e^{j\omega \cdot s_i} H(\omega + \bar{\kappa}_i), \, i \in \{1, 2, \cdots 7\}$$

where the shift vectors s_i are defined as

$$s_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad s_1 = \begin{pmatrix} 1 \\ 0 \\ 1/\sqrt{2} \end{pmatrix},$$

$$s_2 = \begin{pmatrix} 1 \\ 1 \\ \sqrt{2} \end{pmatrix}, \quad s_3 = \begin{pmatrix} 1 \\ 2 \\ 1/\sqrt{2} \end{pmatrix},$$

$$s_4 = \begin{pmatrix} 2 \\ 1 \\ 1/\sqrt{2} \end{pmatrix}, \quad s_5 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix},$$

$$s_6 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad s_7 = \begin{pmatrix} 0 \\ 1 \\ 1/\sqrt{2} \end{pmatrix}.$$

Note that as in the hexagonal case, this choice is not unique.

With the choice of filters given above, the aliasing terms in equation (15) cancel and the remaining LSI system response is

$$\hat{X}(\omega) = \frac{1}{8} X(\omega) \sum_{i=0}^{7} |H(\omega + \bar{\kappa}_i)|^2,$$

independent of the choice of the function $H(\omega)$. The design constraint equation is now

$$\sum_{i=0}^{7} |H(\omega + \bar{\kappa}_i)|^2 = 8.$$

We designed a garnet filter using the same method described in the previous section. To demonstrate its use, we have applied it to an image sequence of a sinusoidal pinwheel rotating in a counterclockwise direction. One frame of the sequence is shown in Fig. 19(a). The squared

responses of the four different motion-selective filters (filters $F_3(\omega)$ through $F_6(\omega)$) are shown in Fig. 19(b)–(e). As expected, the filters extract information within spatio-temporal frequency bands corresponding to motion in the
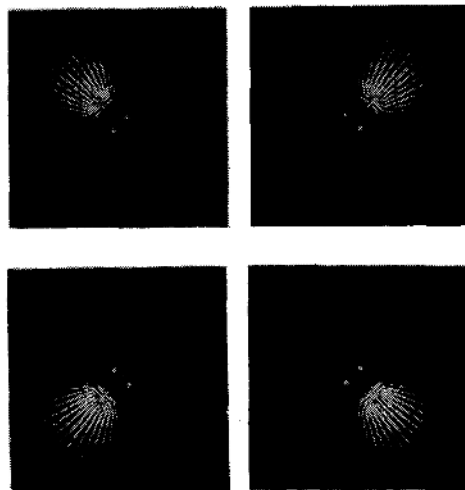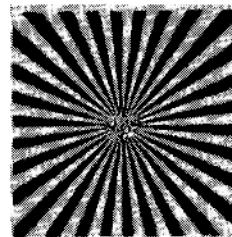

Fig. 19 At the top is one image from a rotating pinwheel image sequence. The four lower images are the squared result of convolving the sequence with four of the "garnet" filters described in the text. Each filter responds preferentially to one direction of motion.

directions down/left, up/left, down/right, and up/right. We believe that the garnet filters will be useful for motion analysis and for data compression of three-dimensional images such as medical imagery and motion sequences.

CONCLUSIONS

In this paper, we have demonstrated that one-dimensional QMF transforms can be extended to multiple dimensions in an orientation-selective manner. The resulting multi-dimensional transform can be applied recursively to create a multi-scale oriented pyramid representation. The filters are relatively compact and computationally efficient, assuming one takes advantage of symmetry. We have demonstrated the usefulness of the transform in the areas of data compression, progressive transmission, orientation analysis, and motion extraction. We believe the transform will be extremely useful for many other machine vision tasks.

APPENDIX: EXAMPLE HEXAGONAL FILTERS

Table 1 A Hexagonal Filter*

```
              j     k     l     l     k     j
        k     g     h     i     h     g     k
     l     h     e     f     f     e     h     l
  l     i     f     c     d     c     f     i     l
k     h     i     d     b     b     d     f     h     k
j  g     e     c     b     a     b     c     e     g     j
k     h     f     d     b     b     d     f     h     k
  l     i     f     c     d     c     f     i     l
     l     h     e     f     f     e     h     l
        k     g     h     i     h     g     k
              j     k     l     l     k     j
```

*The letters refer to the free parameters (see text). Only the low-pass filter is shown. The three highpass filters are formed by modulating and shifting the low-pass.

Table 2 Some Example Hexagonal Filter Coefficient Values

Parameter	3-ring filter	4-ring filter	5-ring filter
a	0.59290695	0.6066799	0.60879886
b	0.32242984	0.3162482	0.31689283
c	−0.016686682	−0.028019974	−0.027267352
d	−0.061579883	−0.0016289932	−0.012790751
e	−0.0020203826	−0.02741341	−0.03874194
f	−0.0038235565	−0.038143888	−0.02383056
g		−0.005958891	0.0008673751
h		0.019682134	0.015554102
i		0.016045252	0.0080001475
j			−0.0009099232
k			−0.0022140248
l			−0.0010486352

REFERENCES

[1] D. Marr, Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. San Francisco, CA: W. H. Freeman and Company, 1982.

[2] J. J. Koenderink, "The structure of images," Biological Cybernetics, vol. 50, pp. 363–370, 1984.

[3] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," GRASP Lab Technical Memo MS-CIS-87-22, University of Pennsylvania, Department of Computer and Information Science, 1987.

[4] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden, "Pyramid methods in image processing," RCA Engineer, vol. 29, no. 6, pp. 33–41, Nov./Dec. 1984.

[5] A. Croisier, D. Esteban, and C. Galand, "Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques," in Int. Conf. on Information Sciences and Systems, (Patras, Aug. 1976), pp. 443–446.

[6] D. Esteban and C. Galand, "Application of quadrature mirror filters to split band voice coding schemes," in Proc. ICASSP, 1977, pp. 191–195.

[7] P. P. Vaidyanathan, "Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M, having the perfect-reconstruction property," IEEE Trans. Acoust. Signal Process., vol. ASSP-35, no. 4, pp. 476–492, Apr. 1987.

[8] M. Vetterli, "Multi-dimensional sub-band coding: some theory and algorithms," Signal Processing, vol. 6, no. 2, pp. 97–112, Feb. 1984.

[9] P. P. Vaidyanathan, "Perfect reconstruction QMF banks for two-dimensional applications," IEEE Trans. Circuits and Systems, vol. CAS 34, no. 8, pp. 976–978, Aug. 1987.

[10] E. Viscito and J. Allebach, "Design of perfect reconstruction multi-dimensional filter banks using cascaded Smith form matrices," in Proc. ISCAS, pp. 831–834, 1988.

[11] J. W. Woods and Sean D. O'Neil, "Subband coding of images," IEEE Trans. Acoust. Signal Process., vol. ASSP-34, no. 5, pp. 1278–1288, Oct. 1986.

[12] E. H. Adelson, E. Simoncelli, and R. Hingorani, "Orthogonal pyramid transforms for image coding," in Proc. SPIE, Oct. 1987.

[13] H. Gharavi and A. Tabatabai, "Application of quadrature mirror filters to the coding of monochrome and color images," in Proc. ICASSP, pp. 32.8.1–32.8.4, 1987.

[14] A. Tran, K-M. Liu, K-H. Tzou, and E. Vogel, "An efficient pyramid image coding system," in Proc. ICASSP, pp. 18.6.1–18.6.4, 1987.

[15] M. Kunt, A. Ikonomopoulos, and M. Kocher, "Second generation image-coding techniques," in Proc. IEEE, pp. 549–574, 1985.

[16] A. B. Watson, "Efficiency of a model human image code," J. Opt. Soc. Am. A, vol. 12, pp. 2401–2417, 1987.

[17] J. G. Daugman and Daniel M. Kammen, "Pure orientation filtering: a scale-invariant image-processing tool for perception research and data compression," Behavior Research Methods, Instruments, & Computers, vol. 18, no. 6, pp. 559–564, 1986.

[18] M. Porat and Y. Zeevi, "The generalized Gabor scheme of image representation in biological and machine vision," IEEE Trans. PAMI, vol. 10, pp. 452–468, 1988.

[19] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," J. Opt. Soc. Am. A, vol. 2, no. 7, pp. 1160–1169, July 1985.

[20] J. P. Crettez and J. C. Simon, "A model for cell receptive fields in the visual striate cortex," Computer Graphics and Image Processing, vol. 20, pp. 299–318, 1982.

[21] A. B. Watson and A. J. Ahumada, "A hexagonal orthogonal-oriented pyramid as a model of image representation in visual cortex," IEEE Trans. Biomedical Engineering, vol. 36, no. 1, pp. 97–106, Jan. 1989.

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

[23] P. P. Vaidyanathan, "Quadrature mirror filter banks, M-band extensions and perfect-reconstruction techniques," IEEE ASSP Magazine, pp. 4–20, July 1987.

[24] M. Vetterli, "A theory of multirate filter banks," IEEE Trans. ASSP, ASSP-35, no. 3, pp. 356–372, Mar. 1987.

[25] E. P. Simoncelli. Orthogonal sub-band image transforms. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, May 1988.

[26] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," in Proc. ICASSP, pp. 291–294, 1980.

[27] V. K. Jain and R. E. Crochiere, "A novel approach to the design of analysis/synthesis filter banks," in Proceedings ICASSP, pp. 5.10–5.10, 1983.

[28] ——, "Quadrature mirror filter design in the time domain," IEEE Trans. ASSP, vol. ASSP-32, no. 2, pp. 353–360, Apr. 1984.

[29] G. Wackersreuther, "On the design of filters for ideal QMF and polyphase filter banks," Arch. Elekt. Ubertrag, vol. 39, pp. 123–130, 1985.

[30] R. M. Mersereau and T. C. Speake, "The processing of periodically sampled multidimensional signals," IEEE Trans. ASSP, vol. ASSP-31, no. 1, pp. 188–194, Feb. 1983.

[31] D. E. Dudgeon and R. M. Mersereau, Multidimensional Digital Signal Processing, Signal Processing Series, Englewood Cliffs, NJ: Prentice-Hall, 1984.

[32] N. S. Jayant and P. Noll, Digital Coding of Waveforms, Signal Processing Series, Englewood Cliffs, NJ: Prentice-Hall, 1984.

[33] W. T. Freeman and E. H. Adelson, The Design and Use of Steerable Filters. Vision Science Technical Report 118, Vision Science Group, Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA, 1989.

Eero P. Simoncelli (Student Member, IEEE) received the B.S. degree in physics, summa cum laude, from Harvard University, Cambridge, MA, in 1984, and the M.S. degree from the Massachusetts Institute of Technology, Cambridge, MA, in 1988.

His research experience includes work in neural networks at Bell Laboratories, done after graduating from Harvard, followed by work in mathematics at Cambridge University, England, as a Knox Fellowship recipient. His M.S. thesis focused on sub-band transform representation of images.

Mr. Simoncelli is a member of Phi Beta Kappa.

Edward H. Adelson (Member, IEEE) received the B.A. degree in physics and philosophy from Yale University, New Haven, CT, in 1974, and the Ph.D. degree in experimental psychology from the University of Michigan, Ann Arbor, in 1979.

He is an Associate Professor of Vision Science at the Media Laboratory and the Department of Brain and Cognitive Science of the Massachusetts Institute of Technology, Cambridge, MA. He has published widely on various topics in human psychophysics, computational vision, and image processing.

Dr. Adelson is a Fellow of the Optical Society of America, and a recipient of the Adolph Lomb Medal.