A Wavelet Multiresolution Analysis for Spatio-Temporal Signals

THOMAS J. BURNS
STEVEN K. ROGERS
MARK E. OXLEY
DENNIS W. RUCK
Air Force Institute of Technology

The wavelet filters of the conventional 3D multiresolution analysis possess homogeneous spatial and temporal frequency characteristics which limits one's ability to match filter frequency characteristics to signal frequency behavior. Also, the conventional 3D multiresolution analysis employs an oct-tree decomposition structure which restricts the analysis of signal details to identical resolutions in space and time. This paper presents a 3D wavelet multiresolution analysis constructed from nonhomogeneous spatial and temporal filters, and an orthogonal sub-band coding scheme that decouples the spatial and temporal decomposition processes.

I. INTRODUCTON

Three-dimensional (i.e., spatio-temporal) signal compression research has seen a steady increase over the last decade in support of the next generation of video communication systems such as video telephones, high definition television, and video teleconferencing systems. Additionally, pattern recognition strategies, which have traditionally focused on spatial information contained in single, static image frames, are now expanding their scope to include temporal information contained in multiple frames of time-sequential imagery [14, 15, 22]. Hierarchical, or "multiresolution," signal processing techniques, which have been used successfully in 2D image compression and feature generation algorithms, are now being extended to 3D signals for these same purposes [3, 4, 21, 26–28]. In particular, wavelet-based multiresolution analyses are gaining popularity because they yield orthonormal building blocks for finite energy functions that are considerably more diverse than the complex exponentials found in conventional Fourier analysis [20]. Additionally, the subband transforms used in discrete multiresolution wavelet decomposition and reconstruction algorithms provide a "fast" method for analyzing and synthesizing signals [17, 25].

Wavelet multiresolution analysis techniques found in the literature have been applied primarily to 1D and 2D signals. These techniques project the signal onto a chain of embedded approximation and detail spaces designed to represent the signal and its details at various levels of resolution [7]. For practical purposes, the projection coefficients are obtained using a discrete subband transform that employs a quadrature mirror filter pair related to the type of wavelet used in the analysis. In Mallat's conventional 2D wavelet multiresolution analysis, the separable 2D approximation spaces are formed from the tensor product of identical 1D approximation spaces [16]. This restriction generates analyzing filters with homogeneous spectral characteristics in 2D frequency space.

When extended to three dimensions, the approximation spaces of the conventional 3D wavelet multiresolution analyses are constructed from the tensor product of three identical 1D approximation spaces [2]. The frequency spectrum of the 3D analyzing filter is therefore identical in space and time, which in turn limits the ability of the filter designer to match the frequency characteristics of the filter with those of the signal. Furthermore, extending the "quad-tree" subband transform of the conventional 2D wavelet multiresolution analysis to three dimensions, produces an "oct-tree" decomposition structure that precludes the possibility of analyzing the signal at different resolutions in space and time [2].

A method is described here for constructing a wavelet multiresolution analysis for 3D signals that

overcomes these problems. First, a general background is provided regarding the wavelet multiresolution analysis in order to familiarize the reader with the notation used here. Next, a 3D wavelet multiresolution analysis is constructed from a separable 3D analyzing, or "scaling," function formed from the product of three nonidentical 1D scaling functions. This yields a much richer set of orthonormal basis vectors with which to represent 3D signals, and it produces filters that can be easily tailored to more closely match the spatial and temporal frequency characteristics of the 3D signal. An unconventional subband transform is then presented which essentially "decouples" the spatial and temporal decomposition processes to yield a wavelet multiresolution tool with independent zoom-in and zoom-out capabilities in space and time. The utility of these properties are demonstrated by applying the nonhomogeneous 3D wavelet multiresolution analysis to various synthetic and real IR image sequences. The final section discusses a method for increasing the orientational selectivity of the decoupled decomposition process.

II. BACKGROUND

An $L_2(\mathbf{R})$ multiresolution analysis consists of a chain of closed, linear "approximation" spaces V_j and a scaling function ϕ which satisfy the following properties for any $f \in L_2(\mathbf{R})$ [5].

1)
$$\cdots V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots .$$

2)
$$\overline{\bigcup_{j \in \mathbf{Z}} V_j} = L_2(\mathbf{R}); \qquad \bigcap_{j \in \mathbf{Z}} V_j = \{0\}.$$

3)
$$f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}; \qquad j \in \mathbf{Z}$$
$$f(x) \in V_j \Rightarrow f\left(x + \frac{n}{2^j}\right) \in V_j; \qquad n \in \mathbf{Z}.$$

4) The set of functions $\{2^{j/2}\phi(2^j x - n) \mid n \in \mathbf{Z}\}$ forms an orthonormal basis for the approximation space V_j.

As presented by Mallat, the purpose of the multiresolution analysis is to create a mathematical framework that facilitates the construction of a wavelet orthonormal basis for the space of all finite energy signals $L_2(\mathbf{R})$. To this end, denote the orthogonal complement of V_j in V_{j+1} by W_j where

$$V_{j+1} = V_j \oplus W_j \qquad (1)$$

and the symbol $\oplus$ indicates the direct sum [16]. W_j is typically referred to as the jth *detail* space, because it captures the difference in signal information between the approximation spaces V_{j+1} and V_j.

Mallat has shown one can create a mother wavelet $\psi(x)$ such that the set of functions $\{2^{j/2}\psi(2^j x - n) \mid n \in \mathbf{Z}\}$ forms an orthonormal basis for W_j. The spaces W_j, where $j \in \mathbf{Z}$, are mutually orthogonal; thus, by the denseness property of the multiresolution analysis, the set of scaled and dilated wavelets $\{2^{j/2}\psi(2^j x - n) \mid j \in \mathbf{Z}, n \in \mathbf{Z}\}$ forms an orthonormal basis for $L_2(\mathbf{R})$. The scaling functions and the mother wavelet are related by the "two-scale" recursion relations

$$\phi(x) = \sum_{n=-\infty}^{\infty} h_n \sqrt{2}\phi(2x - n)$$
$$\psi(x) = \sum_{n=-\infty}^{\infty} g_n \sqrt{2}\phi(2x - n) \qquad (2)$$

where the coefficients h_n and g_n are discussed below.

Approximation and detail signals are created by orthogonally projecting the input signal f onto the appropriate approximation or detail space. Since each space is spanned by an orthonormal basis set, the signal projection onto a given approximation or detail space at, say, the jth resolution, is equivalent to the sequence of projection coefficients obtained by the inner product operations

$$a_{j,n} = \int_{-\infty}^{\infty} f(x) 2^{j/2} \phi(2^j - n) \, dx$$
$$d_{j,n} = \int_{-\infty}^{\infty} f(x) 2^{j/2} \psi(2^j - n) \, dx \qquad (3)$$

where $a_{j,n}$ and $d_{j,n}$ represent the jth approximation and detail coefficients, respectively.

In practice, the approximation and detail projection coefficients associated with V_j and W_j are computed from the approximation coefficients at the next higher scale, V_{j+1}, using a quadrature mirror filter (QMF) pair and a pyramidal subband coding scheme [17]. The impulse responses of the QMF pair are typically denoted by h_n and g_n, where h_n is formed from the inner product between $\phi(u/2)$ and $\phi(u - n)$ and, for this research, $g_n = (-1)^{1-n} h_{1-n}$.

A binary tree structure for implementing Mallat's 1D wavelet multiresolution analysis is shown in Fig. 1(a). The binary tree serves as a "canonical" structure for extending the conventional algorithm to multiple dimensions. Here, the coefficients of the $j + 1$st approximation level are simultaneously decomposed into the jth detail and approximation coefficients using the low-pass and high-pass impulse responses $h(n)$ and $g(n)$. The regions of support in frequency space of the resulting approximation and detail signals are shown in Fig. 1(b). By repeatedly convolving each approximation signal with $h(n)$ and $g(n)$ and decimating the outputs by a factor of two, the signal is decomposed into frequency bands whose bandwidths and center frequencies vary by octaves. In the signal processing literature, the set of filters generated by multiple stages of the the pyramidal decomposition
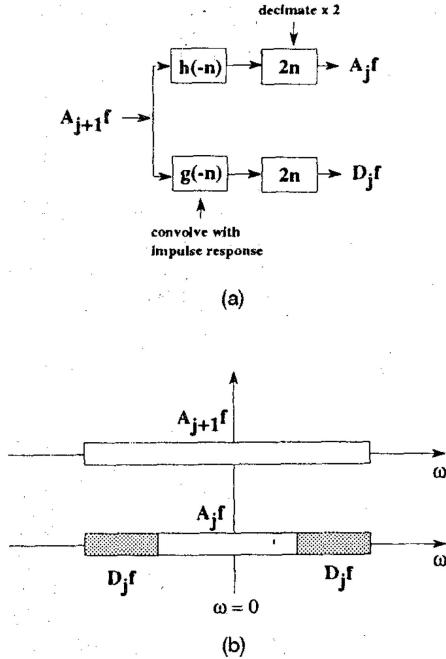
Fig. 1. (a) 1D subband transform for decomposing coefficients of $j + 1$st approximation level into coefficients of jth detail and approximation levels. (b) Regions of support along frequency axis of approximation and detail signals.

algorithm is referred to as a two-channel paraunitary QMF filter bank [25].

One can also construct a separable orthonormal wavelet basis set for $L_2(\mathbf{R}^2)$ (the space containing all finite energy 2D signals) from the chain of 2D multiresolution approximation spaces, $\{\mathbf{V}_j \mid j \in \mathbf{Z}\}$, where $\mathbf{V}_j$ is defined by [7, 18]

$$\mathbf{V}_j = V_j^x \otimes V_j^y$$
$$= \overline{\text{Span}\{F(x,y) = f(x)g(y) \mid f \in V_j^x,\ g \in V_j^y\}}$$
$$(4)$$

$$F(x,y) \in \mathbf{V}_j \Leftrightarrow F(2x, 2y) \in \mathbf{V}_{j+1} \qquad (5)$$

and where V_j^x and V_j^y are identical 1D approximation spaces (i.e., they are spanned by the same scaling function). Here, the 2D scaling function for $\mathbf{V}_j$ is formed from the product of both identical 1D scaling functions, and the wavelet orthonormal basis for the orthogonal complement $\mathbf{W}_j$ is given by three wavelets

$$\Psi_j^1(x,y) = 2^j \phi(2^j x)\psi(2^j y)$$
$$\Psi_j^2(x,y) = 2^j \psi(2^j x)\phi(2^j y) \qquad (6)$$
$$\Psi_j^3(x,y) = 2^j \psi(2^j x)\psi(2^j y).$$

The family of wavelets

$$\{\Psi_j^p(x-m, y-n) \mid j \in \mathbf{Z};\ (m,n) \in \mathbf{Z}^2;\ p = 1,2,3\} \qquad (7)$$

then forms an orthonormal basis set for $L_2(\mathbf{R}^2)$. Through a straightforward extension of the 1D

binary tree structure, one obtains the 2D quad-tree wavelet multiresolution decomposition algorithm shown in Fig. 2(a). Here $A_j f$ and $D_j^n f$, $n = 1, 2, 3$ denote the projection of the $L_2(\mathbf{R}^2)$ image f onto the approximation space $\mathbf{V}_j$ and detail spaces $\mathbf{W}_j^1$, $\mathbf{W}_j^2$, and $\mathbf{W}_j^3$ spanned, respectively, by the wavelets $\{\Psi_j^p \mid p = 1,2,3\}$ in (6) [17]. Fig. 2(b) shows the idealized frequency supports of the separable approximation and detail filters used to decompose the 2D image approximation $A_{j+1} f$ into the approximation $A_j f$ and the details D_j^1, D_j^2, and D_j^3. Here, the wavelet filters Ψ_j^1, Ψ_j^2, and Ψ_j^3, respectively, capture the horizontal, vertical, and diagonal details in the image at the jth resolution.

Because the homogeneous 2D approximation spaces $\mathbf{V}_j$ are constructed from the tensor product of two identical 1D approximation spaces, the same discrete filters, $h(n)$ and $g(n)$, are convolved with both the rows and columns of the input image approximation in Fig. 2(b). When extended to three dimensions, the identical filter pairs are now convolved with the rows, columns, and *frames* of the input image approximation sequence [2]. However, spatio-temporal image sequences clearly do not possess homogeneous spatial and temporal frequency characteristics. For example, a small object moving slowly across an image plane will have a broad spatial frequency spectrum and a narrow temporal frequency spectrum. Therefore, the following section shows one can construct a *nonhomogeneous* 3D multiresolution wavelet analysis that allows the filter designer to independently control the spatial and temporal frequency characteristics of the filter while maintaining the orthogonality of the resulting wavelet representation. Proofs of the Propositions put forth in this section are contained in Appendix A.

III. A NONHOMOGENEOUS 3D MULTIRESOLUTION ANALYSIS AND OCT-TREE DECOMPOSITION ALGORITHM

The previous section discussed how one constructs an orthonormal wavelet basis for $L_2(\mathbf{R})$ and $L_2(\mathbf{R}^2)$ signals through the mathematical framework of a homogeneous wavelet multiresolution analysis. This section describes how to construct an orthonormal wavelet basis for $L_2(\mathbf{R}^3)$ signals using a nonhomogeneous multiresolution analysis. The nonhomogeneous multiresolution analysis produces wavelets with spatial and temporal frequency characteristics that can be independently varied to more closely match the frequency behavior of the signal under analysis. An oct-tree decomposition algorithm is then described for computing the discrete coefficients associated with the orthonormal wavelet series representation of an $L_2(\mathbf{R}^3)$ signal.

In order to construct a nonhomogeneous 3D multiresolution analysis, one must construct a chain
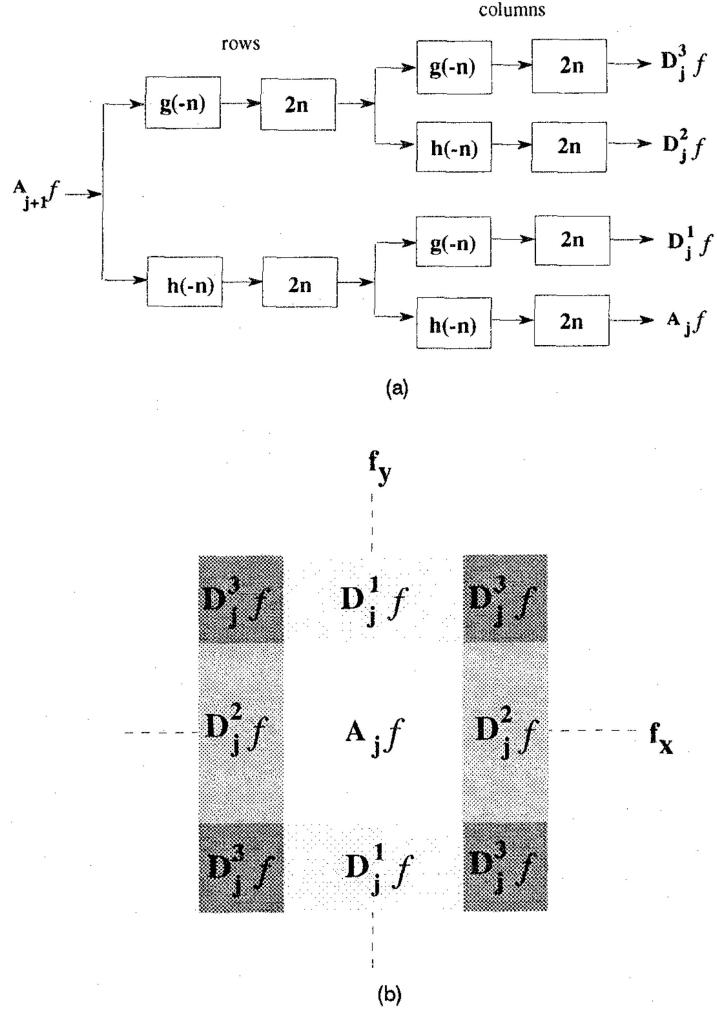
Fig. 2.   (a) Mallat's 2D discrete multiresolution decomposition algorithm. (b) Idealized frequency representation of the 2D decomposition of the approximation image $A_{j+1}f$ into A_jf and the detail images $D_j^p f$; $p = 1,2,3$ [17].

of approximation spaces and a scaling function that meet the properties outlined in the previous section. To this end, begin by selecting a scaling function ϕ such that the set of functions $\{2^{j/2}\phi(2^j \cdot -n) \mid n \in \mathbf{Z}\}$ forms an orthonormal basis for the multiresolution approximation V_j of $L_2(\mathbf{R})$. Next, let $\tilde{\phi}$ be a different scaling function such that $\{2^{j/2}\tilde{\phi}(2^j \cdot -n) \mid n \in \mathbf{Z}\}$ forms an orthonormal basis for the multiresolution approximation $\tilde{V}_j$ of $L_2(\mathbf{R})$. Now define the separable, closed, linear subspaces of $L_2(\mathbf{R}^3)$ by

$$\mathbf{V}_j = V_j^x \otimes V_j^y \otimes \tilde{V}_j^t$$

$$= \overline{\text{Span}\{F(x,y,t) = f(x)g(y)h(t) \mid f \in V_j^x, \ g \in V_j^y \text{ and } h \in \tilde{V}_j^t\}}.$$

$$(8)$$

Given the pair of 1D scaling functions, ϕ and $\tilde{\phi}$, and the above definition of the approximation space $\mathbf{V}_j$, Proposition 1 shows one can construct a separable 3D scaling function such that the set comprised of all its integer translations forms an orthonormal basis for the nonhomogeneous approximation space $\mathbf{V}_j$.

PROPOSITION 1   *For each $j \in \mathbf{Z}$, the set of functions $\{2^{3j/2}\phi(2^j x - l)\phi(2^j y - m)\tilde{\phi}(2^j t - n) \mid (l,m,n) \in \mathbf{Z}^3\}$ forms an orthonormal basis for $\mathbf{V}_j$.*

Proposition 1 shows that the nonhomogeneous approximation space $\mathbf{V}_j$ is spanned by integer translations of a separable scaling function formed from the product of three nonidentical scaling functions. Proposition 2 proves that the scaling function and approximation space generate a multiresolution analysis.

PROPOSITION 2   *The family of closed, linear spaces, $\{\mathbf{V}_j \mid j \in \mathbf{Z}\}$, forms a multiresolution analysis in $L_2(\mathbf{R}^3)$.*

In the 3D multiresolution analysis, approximations of the 3D signal at the jth and $(j + 1)$st resolutions in space and time are obtained by orthogonally projecting the signal, respectively, onto the spaces $\mathbf{V}_j$ and $\mathbf{V}_{j+1}$. The spatial and temporal details that comprise the difference in information between these two approximations are then contained in the orthogonal complement of $\mathbf{V}_j$ in $\mathbf{V}_{j+1}$. Denoting this

complementary space by the symbol $\mathbf{W}_j$, Proposition 3 shows one can create an orthonormal basis for $\mathbf{W}_j$ (and for $L_2(\mathbf{R}^3)$) from seven sets of scaled and translated "wavelets."

PROPOSITION 3 *Let ψ and $\tilde{\psi}$ be the one-dimensional wavelets, respectively, generated by the scaling functions ϕ and $\tilde{\phi}$. Then the seven "wavelets"*

$$\Psi_j^1(x,y,t) = 2^{3j/2}\phi(2^j x)\phi(2^j y)\tilde{\psi}(2^j t)$$

$$\Psi_j^2(x,y,t) = 2^{3j/2}\phi(2^j x)\psi(2^j y)\tilde{\phi}(2^j t)$$

$$\Psi_j^3(x,y,t) = 2^{3j/2}\phi(2^j x)\psi(2^j y)\tilde{\psi}(2^j t)$$

$$\Psi_j^4(x,y,t) = 2^{3j/2}\psi(2^j x)\phi(2^j y)\tilde{\phi}(2^j t) \qquad (9)$$

$$\Psi_j^5(x,y,t) = 2^{3j/2}\psi(2^j x)\phi(2^j y)\tilde{\psi}(2^j t)$$

$$\Psi_j^6(x,y,t) = 2^{3j/2}\psi(2^j x)\psi(2^j y)\tilde{\phi}(2^j t)$$

$$\Psi_j^7(x,y,t) = 2^{3j/2}\psi(2^j x)\psi(2^j y)\tilde{\psi}(2^j t)$$

are such that for each $j \in \mathbf{Z}$, $\{\Psi_j^p(x-l, y-m, t-n) \mid (l,m,n) \in \mathbf{Z}^3; p = 1, 2, \ldots, 7\}$ forms an orthonormal basis for $\mathbf{W}_j$ and $\{\Psi_j^p(x-l, y-m, t-n) \mid j \in \mathbf{Z}; (l,m,n) \in \mathbf{Z}^3; p = 1, 2, \ldots, 7\}$ forms an orthonormal basis for $L_2(\mathbf{R}^3)$.

The spatial and temporal details between $\mathbf{V}_{j+1}$ and $\mathbf{V}_j$ contained in $\mathbf{W}_j$ are obtained by orthogonally projecting the 3D signal onto the orthonormal basis set described in Proposition 3. Since this process is carried out on a digital computer, the projection operation is typically represented in discrete form by the coefficients generated by computing the inner product of the signal with all possible integer translations of the functions in (9). Following Mallat, it can be shown that the approximation and detail coefficients at the jth resolution can be computed using the following discrete convolution operations [2, 17]

$$a_{j;l,m,n} = [a_{j+1;p,q,r} * \overline{h}(p) * \overline{h}(q) * \overline{\tilde{h}}(r)](2l, 2m, 2n)$$

$$d_{j;l,m,n}^1 = [a_{j+1;p,q,r} * \overline{h}(p) * \overline{h}(q) * \overline{\tilde{g}}(r)](2l, 2m, 2n)$$

$$d_{j;l,m,n}^2 = [a_{j+1;p,q,r} * \overline{h}(p) * \overline{g}(q) * \overline{\tilde{h}}(r)](2l, 2m, 2n)$$

$$d_{j;l,m,n}^3 = [a_{j+1;p,q,r} * \overline{h}(p) * \overline{g}(q) * \overline{\tilde{g}}(r)](2l, 2m, 2n)$$

$$d_{j;l,m,n}^4 = [a_{j+1;p,q,r} * \overline{g}(p) * \overline{h}(q) * \overline{\tilde{h}}(r)](2l, 2m, 2n) \qquad (10)$$

$$d_{j;l,m,n}^5 = [a_{j+1;p,q,r} * \overline{g}(p) * \overline{h}(q) * \overline{\tilde{g}}(r)](2l, 2m, 2n)$$

$$d_{j;l,m,n}^6 = [a_{j+1;p,q,r} * \overline{g}(p) * \overline{g}(q) * \overline{\tilde{h}}(r)](2l, 2m, 2n)$$

$$d_{j;l,m,n}^7 = [a_{j+1;p,q,r} * \overline{g}(p) * \overline{g}(q) * \overline{\tilde{g}}(r)](2l, 2m, 2n)$$

where $*$ indicates the discrete convolution operator, $a_{j;l,m,n}$ is a 3D approximation coefficient at the location (l,m,n) in a cubic sampling lattice, $\{d_{j;l,m,n}^p \mid p = 1, \ldots, 7\}$ are the 3D detail coefficients associated with the seven wavelets in Proposition 3, $\overline{h}(n) = h(-n)$, $\overline{\tilde{h}}(n) = \tilde{h}(-n)$, and the two QMF pairs $(h(n), g(n))$, $(\tilde{h}(n), \tilde{g}(n))$ are obtained from the relationships

$$h(n) \equiv \int_{-\infty}^{\infty} \phi\left(\frac{x}{2}\right)\phi(x-n)dx \qquad (11)$$

$$\tilde{h}(n) \equiv \int_{-\infty}^{\infty} \tilde{\phi}\left(\frac{t}{2}\right)\tilde{\phi}(t-n)dt \qquad (12)$$

and

$$g(n) = (-1)^{1-n}h(1-n)$$
$$\tilde{g}(n) = (-1)^{1-n}\tilde{h}(1-n). \qquad (13)$$

Equation (10) indicates the approximation and detail coefficients at the jth resolution are computed by discretely convolving the coefficients at the next higher resolution level with the separable, nonhomogeneous 3D impulse responses formed from eight different combinations of the spatial and temporal QMF pairs and decimating by a factor of two in each dimension. These convolution operations then form the eight branches of the nonhomogeneous oct-tree decomposition structure shown in Fig. 3. The oct-tree decomposition structure is referred to here as "conventional" because it is constructed by appending the canonical branch structure of the conventional 1D multiresolution analysis on each of the four branches of the 2D quad-tree structure.

To help visualize the signal processing properties of the 3D wavelet multiresolution analysis, consider the idealized frequency supports shown in Fig. 4 of the filters generated by Fourier transforming the scaling function and the seven "wavelets" described in Propositions 1 and 3. The frequency supports now form a volume in spatio-temporal frequency space, where the inner core of the volume represents the frequency support of the scaling function for the approximation space $\mathbf{V}_j$. The surrounding layers represent the frequency support of the functions $\Psi_j^1 - \Psi_j^7$ that span each of the seven detail spaces. Taken together, the entire volume comprises the frequency support of the scaling function associated with the approximation space $\mathbf{V}_{j+1}$. Therefore, viewed from a filtering perspective, the nonhomogeneous 3D wavelet decomposition decomposes the signal into an independent set of *spatio-temporally* oriented frequency channels, where the spatial and temporal frequency characteristics of the 3D wavelet filters (e.g., size of passband, transition region roll-off, etc.) can be easily varied by mixing and matching a variety of wavelet QMF pairs provided in the literature [5, 7, 17].

The nonhomogeneous 3D wavelet multiresolution analysis provides a way to independently control the spatial and temporal frequency characteristics of the
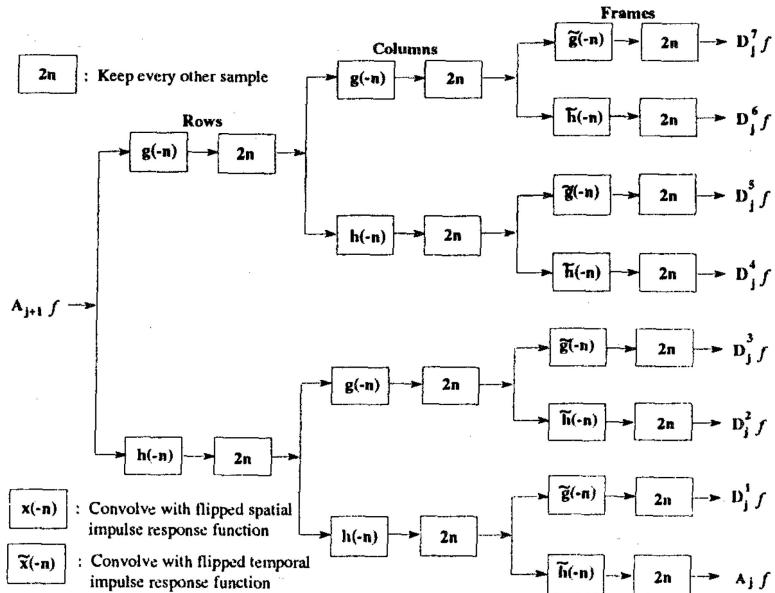
Fig. 3. Oct-tree sub-band transform used to decompose the $j + 1$st approximation coefficients into the jth approximation and detail coefficients in a conventional, non-homogeneous $L_2(\mathbf{R}^3)$ wavelet multiresolution analysis.


(a)

(b)
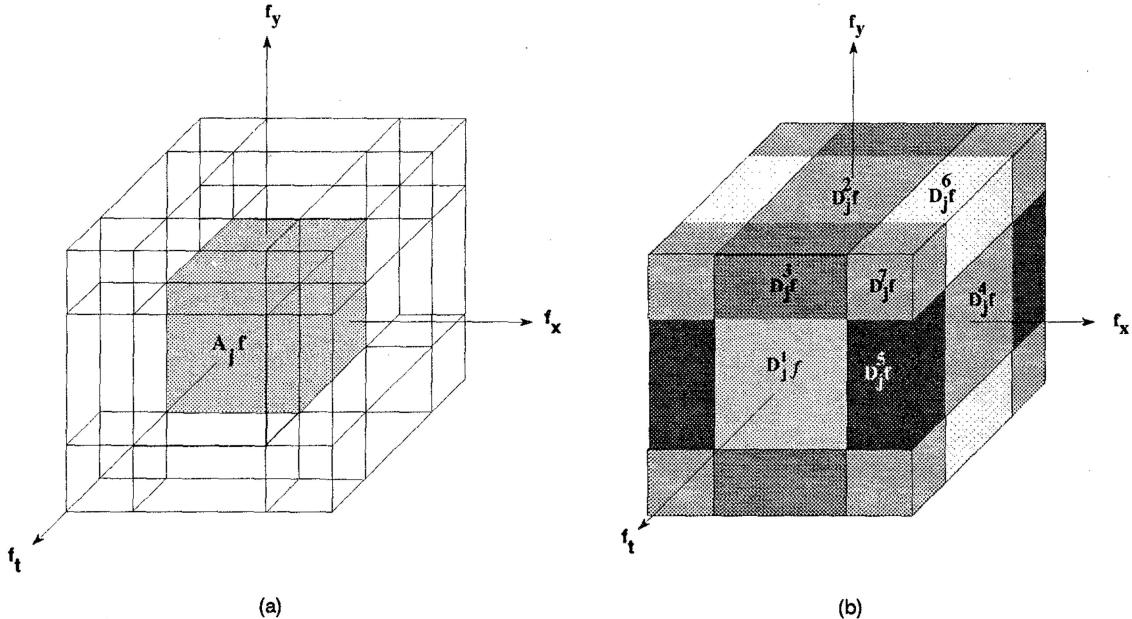
Fig. 4. (a) Idealized frequency support of the approximation signal $A_j f$ generated by a 3D multiresolution decomposition. (b) Idealized frequency supports of the seven surrounding detail signals. $D_j^p f$, $p = 1, 2, \ldots, 7$ represents the projection of the signal onto the detail space spanned by integer translations of the pth wavelet in (9).

wavelet filter while maintaining the orthogonality of the wavelet basis functions. However, the oct-tree decomposition structure used to compute the detail coefficients of the multiresolution analysis is quite restrictive in that it limits the analysis of signal details to the same resolution in space in time. That is, at each stage of the decomposition process the rows, columns, *and* frames of the next higher approximation sequence are each filtered and downsampled by a factor of two. The following section discusses why this

restriction poses a problem for spatio-temporal signal analysis.

IV. A DRAWBACK TO USING THE OCT-TREE DECOMPOSITION ALGORITHM FOR SPATIO-TEMPORAL ANALYSIS

Consider the spatio-temporal cosine function

$$f(x, y, t) = \cos(ax + by + ct). \qquad (14)$$
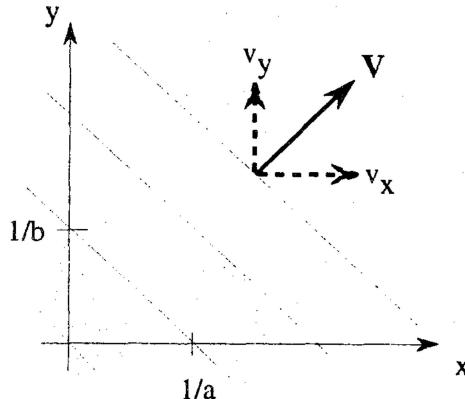
Fig. 5. Lines of constant phase associated with the traveling $\cos f(x,y,t) = \cos(ax + by + ct)$.

If the time-dependent cosine is evaluated at $t = 0$, then, following Goodman, it can be represented by a family of parallel lines of constant phase as shown in Fig. 5 [11]. Here, each line represents a locus of points along which $\cos(ax + by) = 1$ or, equivalently, $ax + by = 2k\pi$, $k = 0,1,2,\ldots$. The velocity of the traveling wave is depicted by the vector $\mathbf{V}$ drawn perpendicular to the lines of constant phase in Fig. 5(b). Assuming the x and y velocity components of $\mathbf{V}$ are given by v_x and v_y, the traveling sinusoid can be expressed as

$$f(x,y,t) = \cos(a(x - v_x t) + b(y - v_y t)) \quad (15)$$

where $c = -av_x - bv_y$. Combining the ratio $v_y/v_x = b/a$ obtained from like triangles in Fig. 5 with (14) and (15), yields the following relationship between the velocity components of the wave and the spatio-temporal frequencies (a,b,c)

$$\begin{aligned} v_x &= -\frac{ac}{a^2 + b^2} \\ v_y &= -\frac{bc}{a^2 + b^2}. \end{aligned} \quad (16)$$

Thus, the spatial frequency pair (a,b) define the pitch, orientation, and direction of the traveling wave, while its speed is directly proportional to the temporal frequency c [10].

Now consider a simple image sequence constructed from the superposition of several sinusoids with identical spatial frequency components, say, a_0 and b_0, traveling with different velocities (i.e., different temporal frequency components). In order to determine the spatio-temporal frequency content of the signal, a filter bank is required that partitions the temporal frequency axis into multiple temporal frequency bands (or resolutions) for a band of spatial frequencies surrounding a_0, b_0. Unfortunately, as shown previously in Fig. 4, the filter bank generated by the oct-tree decomposition structure only produces two temporal frequency bands (low-pass and band-pass) for each spatial frequency band at each resolution. Furthermore, the spatial and temporal bandwidths of the analysis filters both decrease *equally* by a

factor of two from one stage of the decomposition to the next, yielding a filter bank that precludes analysis of the image sequence at different spatial and temporal resolutions. Thus, it is not possible with the conventional structure to construct a filter that discriminates moving objects with dissimilar spatial and temporal frequency characteristics such as large, fast objects (i.e., objects with high temporal frequency and low spatial frequency content), and small, slow objects (low temporal frequency and high spatial frequency content). In order to correct this problem, the next section presents an unconventional subband transform that decouples the spatial and temporal decomposition processes of the conventional multiresolution analysis. The transform is based on a special case of Coifman's wavelet packet theory as described by Daubechies [6, 7].

V. DECOUPLING THE CONVENTIONAL NONHOMOGENEOUS WAVELET MULTIRESOLUTION ANALYSIS

The nonhomogeneous 3D wavelet multiresolution analysis described in the previous section yields a rich set of orthonormal bases for $L_2(\mathbf{R}^3)$ signals that are constructed by combining different spatial and temporal scaling functions. However, the conventional algorithm used to compute the coefficients of the projection onto a spatio-temporal detail space restricts the analysis of the signal to the same resolution in space and time. In order to correct this problem, this section describes an unconventional multiresolution analysis that allows one to analyze a spatio-temporal signal across multiple temporal resolutions for a fixed spatial resolution. Additionally, the wavelet basis set for each of the detail spaces produced by this multiresolution analysis remains orthogonal over all resolutions in space and time, so that, as before, any $L_2(\mathbf{R}^3)$ signal can be uniquely represented by the coefficients obtained under the multiresolution analysis.

The unconventional multiresolution analysis is based on the construction of an orthonormal basis for the decoupled spatio-temporal approximation space $\mathbf{V}_{j,k}$, where

$$\begin{aligned} \mathbf{V}_{j,k} &= V_j^x \otimes V_j^y \otimes \tilde{V}_k^t \\ &= \overline{\mathrm{Span}\{F(x,y,t) = f(x)g(y)h(t) \mid f \in V_j^x,\ g \in V_j^y \text{ and } h \in \tilde{V}_k^t\}} \end{aligned}$$
$$(17)$$

Here, j represents spatial resolution and k represents temporal resolution where j is not necessarily equal to k. The corresponding orthonormal basis for $\mathbf{V}_{j,k}$ is then given by the set of functions $\{2^{j+k/2}\phi(2^j x - l) \cdot \phi(2^j y - m)\tilde{\phi}(2^k t - n) \mid (l,m,n) \in \mathbf{Z}^3\}$ where, as before, ϕ and $\tilde{\phi}$ are different scaling functions.

A straightforward consequence of the definition of the decoupled approximation space is that $\mathbf{V}_{j,k}$ is
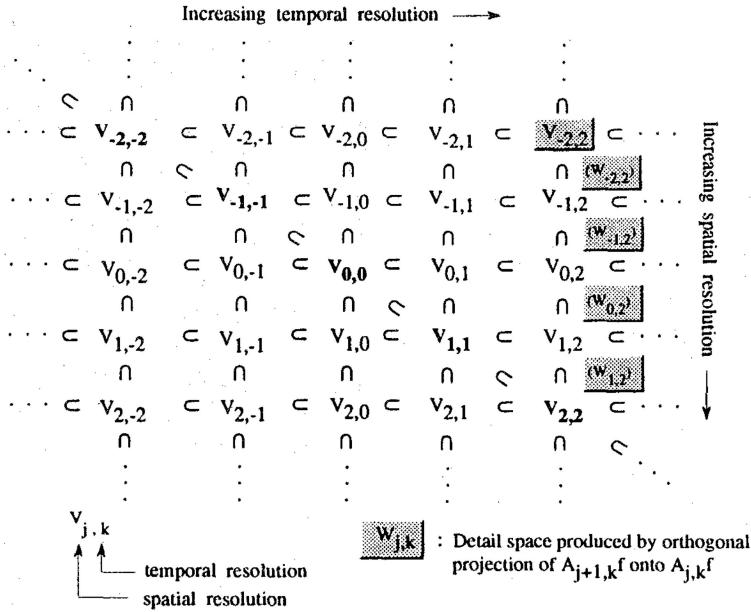
Fig. 6. Lattice of approximation spaces formed by further decomposing conventional multiresolution approximation spaces (represented by diagonal $j = k$) along spatial (vertical) and temporal (horizontal) lines.

contained in $\mathbf{V}_{j',k'}$ if and only if $j \leq j'$ and $k \leq k'$. This fact is illustrated by the lattice of spaces shown in Fig. 6. Here, the chain of spaces comprising the conventional nonhomogeneous 3D multiresolution analysis lies along the diagonal $j = k$. The remaining approximation spaces are created by independently decomposing the conventional spaces $\mathbf{V}_{j,j}$ along spatial (vertical) and temporal (horizontal) lines.

The key elements of the approximation lattice are the "spatial" detail spaces highlighted in gray. These detail spaces, which represent the orthogonal complement of $\mathbf{V}_{j,k}$ in $\mathbf{V}_{j+1,k}$ such that

$$\mathbf{V}_{j+1,k} = \mathbf{V}_{j,k} \oplus \mathbf{W}_{j,k} \qquad (18)$$

contain the horizontal, vertical, and diagonal *spatial* details between the approximation spaces $\mathbf{V}_{j+1,k}$ and $\mathbf{V}_{j,k}$ for a fixed temporal resolution k. Each of the highlighted "spatial" detail spaces are orthogonal and each is spanned by the set of orthogonal functions $\{\Psi_{j,k}^p(x-l,y-m,t-n) \mid (l,m,n) \in \mathbf{Z}^3; \, p=1,2,3\}$ where

$$\Psi_{j,k}^1(x,y,t) = 2^{j+k/2}\phi(2^j x)\psi(2^j y)\tilde{\phi}(2^k t)$$

$$\Psi_{j,k}^2(x,y,t) = 2^{j+k/2}\psi(2^j x)\phi(2^j y)\tilde{\phi}(2^k t) \qquad (19)$$

$$\Psi_{j,k}^3(x,y,t) = 2^{j+k/2}\psi(2^j x)\psi(2^j y)\tilde{\phi}(2^k t).$$

Thus, the spatial detail space $\mathbf{W}_{j,k}$ is comprised of three orthogonal spaces $\mathbf{W}_{j,k}^1$, $\mathbf{W}_{j,k}^2$, and $\mathbf{W}_{j,k}^3$ spanned, respectively by integer translations in space and time of the three wavelets in (19). Fig. 7 shows the idealized regions of support in the positive temporal half of 3D frequency space of these spatio-temporal wavelets, where the frequency spectrum of the approximation signal $\mathbf{V}_{j+1,k}$ has been subdivided by

the spatial decomposition stage into the approximation spectrum associated with $\mathbf{V}_{j,k}$ and the three spatial detail spectrums captured by the filtering properties of the wavelets. In essence, the three filters extract the horizontal, vertical, and diagonal spatial details in the signal at the jth spatial resolution for the kth approximation level in time. In the next stage of the unconventional multiresolution analysis, each spatial detail/temporal approximation space $\mathbf{W}_{j,k}$ is now recursively decomposed in time to yield a sequence of orthogonal spatio-temporal detail spaces that capture the temporal details in an image sequence at multiple resolutions in time for a fixed spatial resolution j.

PROPOSITION 4   *Let $\mathbf{W}_{j,k}^p$ ($p = 1,2,3$) represent the space spanned by integer translations in space and time of the function $\Psi_{j,k}^p(x,y,t) = \Psi_j^p(x,y)2^{k/2}\tilde{\phi}(2^k t)$ where,*

$$\Psi_j^1(x,y) = 2^j\phi(2^j x)\psi(2^j y)$$

$$\Psi_j^2(x,y) = 2^j\psi(2^j x)\phi(2^j y) \qquad (20)$$

$$\Psi_j^3(x,y) = 2^j\psi(2^j x)\psi(2^j y).$$

Define the functions,

$$\Psi_{j,k}^{p1}(x,y,t) = \Psi_j^p(x,y)\sum_n 2^{k/2}\tilde{h}_n\tilde{\phi}(2^k t - n)$$

$$\Psi_{j,k}^{p2}(x,y,t) = \Psi_j^p(x,y)\sum_n 2^{k/2}\tilde{g}_n\tilde{\phi}(2^k t - n) \qquad (21)$$

then the set of functions $\{\Psi_{j,k}^{p1}(x-l,y-m,t-2n),$ $\Psi_{j,k}^{p2}(x-l,y-m,t-2n) \mid (l,m,n) \in \mathbf{Z}^3\}$ forms an orthonormal basis for

$$\mathbf{W}_{j,k}^p = \overline{\mathrm{Span}\{\Psi_{j,k}^p(x-l,y-m,t-n) \mid (l,m,n) \in \mathbf{Z}^3\}}.$$
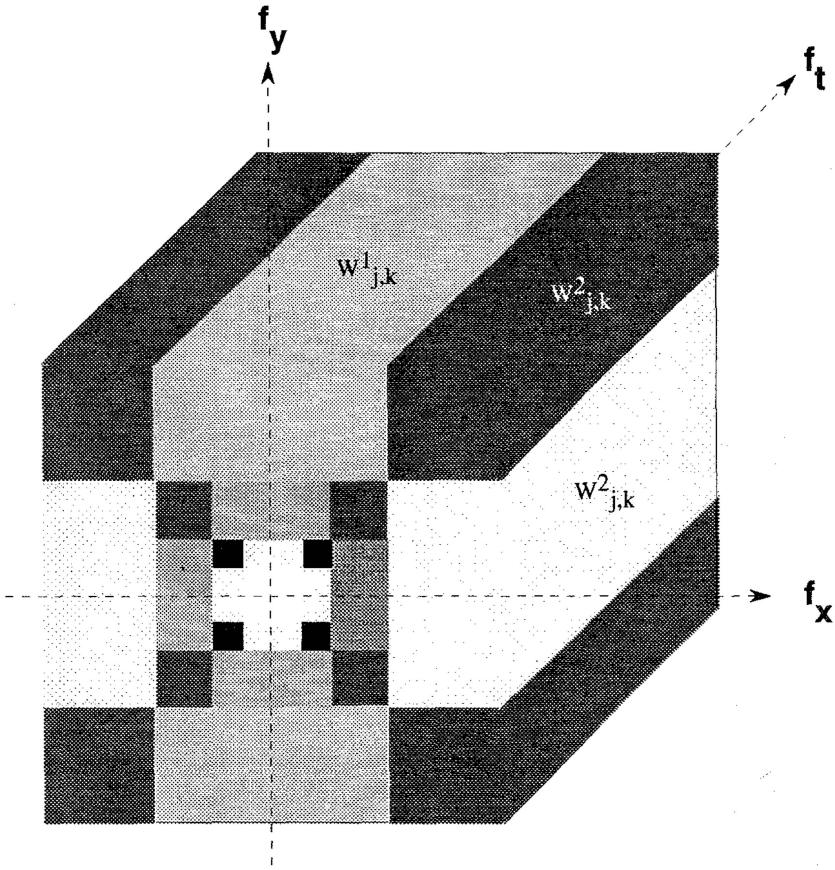
Fig. 7. Idealized supporting regions in frequency space of the three spatio-temporal wavelets in (19) for three resolution levels in space and a constant resolution in time. $W_{j,k}^1$, $W_{j,k}^2$, and $W_{j,k}^3$ correspond to the spatial detail/temporal approximation spaces spanned respectively by integer translations in space and time of the three wavelets.

Proposition 4 constructs an orthonormal basis for a given spatial detail space $\mathbf{W}_{j,k}^p$ by decomposing the basis set for $\mathbf{W}_{j,k}^p$ independently in time using the temporal QMF pair $\tilde{h}, \tilde{g}$. Applying "Daubechies Lemma" (see Appendix A), it is not difficult to show that the set of functions $\{\Psi_{j,k}^{p1}(x-l, y-m, t-2n) \mid (l,m,n) \in \mathbf{Z}^3\}$ then forms an orthonormal basis for the space $\mathbf{W}_{j,k-1}^p$, and the set of functions $\{\Psi_{j,k}^{p2}(x-l, y-m, t-2n) \mid (l,m,n) \in \mathbf{Z}^3\}$ forms an orthonormal basis for the spatio-temporal detail space contained between $\mathbf{W}_{j,k}^p$ and $\mathbf{W}_{j,k-1}^p$. By recursively applying the series operations in Proposition 4 to each approximation space $\mathbf{W}_{j,k}^p$, $\mathbf{W}_{j,k-1}^p$, etc., one can now construct a sequence of spatio-temporal detail spaces that are orthogonal across all temporal resolutions for a fixed spatial resolution j. Fig. 8 shows the idealized regions of support in 3D frequency space of the wavelets created by the decoupled spatio-temporal decomposition process. Unlike the conventional 3D wavelet multiresolution analysis, this procedure subdivides 3D frequency space into multiple temporal frequency bands for each of the horizontal, vertical and diagonal spatial frequency bands at the jth spatial resolution.

Because each spatial detail space $\mathbf{W}_{j,k}$ is orthogonal across all spatial resolutions, Proposition 4 ensures that the resulting *array* of spatio-temporal details obtained by decomposing $\mathbf{W}_{j,k}^p$ for all j and k remains orthogonal across all resolutions in space and time. Thus, the unconventional multiresolution analysis generates a nonuniform filter bank in 3D frequency space that provides the ability to analyze horizontal, vertical, and spatial details independently at any dyadic resolution in space and time, while maintaining the orthogonality of the wavelet representation. The following section describes a fast wavelet transform for computing the coefficients obtained by projecting a 3D image sequence onto the spatio-temporal detail and approximation spaces produced by this process.

VI. A FAST WAVELET TRANSFORM FOR THE UNCONVENTIONAL WAVELET MULTIRESOLUTION ANALYSIS

The previous section ensures that the spatial and temporal decomposition processes in the conventional multiresolution analysis can be decoupled to generate multiple temporal resolutions of a 3D signal for a fixed spatial resolution. Furthermore, the spaces containing
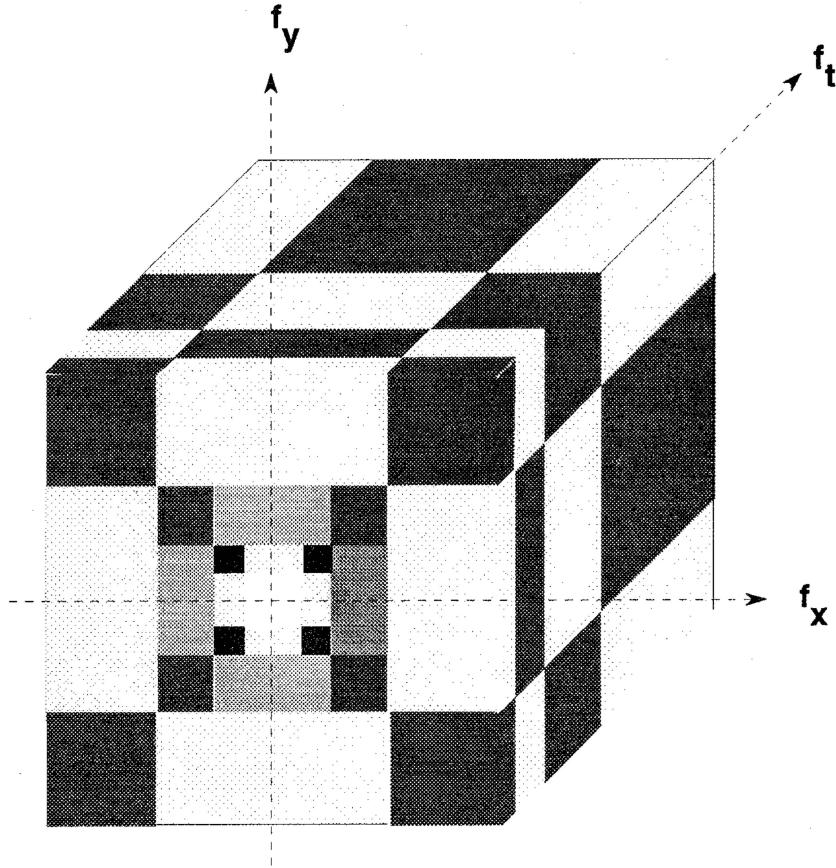
Fig. 8. Idealized supporting regions in 3D frequency space of the wavelets that span the spatio-temporal detail spaces produced by the decoupled wavelet multiresolution analysis for three resolution levels in space.
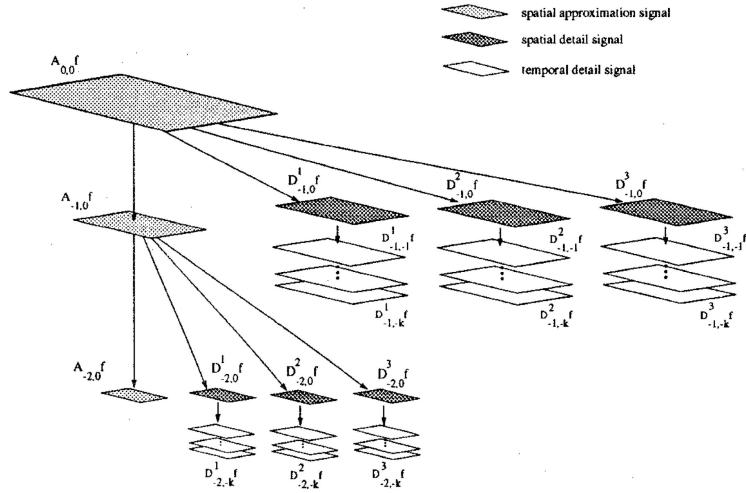

Fig. 9. A visualization of the fast wavelet transform for the unconventional 3D wavelet multiresolution analysis.

the temporal detail signals are orthogonal across all spatial and temporal resolution levels. This section describes an $\mathcal{O}(N^3)$ subband decomposition algorithm that generates the coefficients obtained by orthogonally projecting a 3D signal onto each of these orthogonal detail spaces.

Following Mallat, begin by assuming $A_{0,0}f$ (i.e., the discrete projection of the original signal onto

the space $\mathbf{V}_{0,0}$) represents a sampled 3D image sequence [17]. Additionally, let $D^p_{-1,0}f$ represent the discrete projection of the signal onto the spatial detail spaces $\mathbf{W}^p_{-1,0}$ where it is understood that $p = 1, 2, 3$. Now consider the visualization of the discrete decomposition algorithm shown in Fig. 9.

In the first stage of the decomposition algorithm, $A_{0,0}f$ is decomposed *spatially* into the approximation
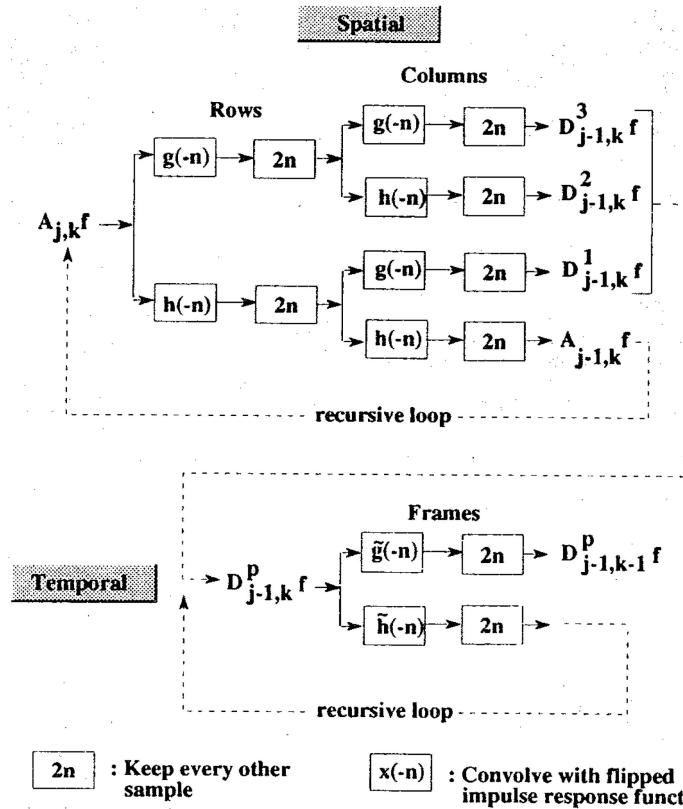
Fig. 10. Spatial, and temporal decomposition algorithms for the unconventional 3D multiresolution wavelet analysis. The decomposition is shown for arbitrary spatial and temporal resolutions levels j and k, respectively.

and detail sequences $A_{-1,0}f$ and $D_{-1,0}^{p}f$ by convolving the rows and columns of each frame in $A_{0,0}f$ with flipped versions of the spatial filters h and g and decimating the spatial dimensions by a factor of two. This process is illustrated in Fig. 10 for arbitrary spatial and temporal resolution levels j and k. The spatial algorithm is then applied recursively to each subsequent spatial approximation sequence, $A_{-j,0}$; $j = 1,2,3\ldots$, to generate a multiresolution ladder of sequences which captures the spatial details between successively smaller spatial resolutions for the temporal resolution $k = 0$. The spatial approximation signals, $A_{-j,0}$; $j = 1,2,3\ldots$, produced by this process are represented by the lightly shaded planes in Fig. 9. The darker planes represent the spatial detail signals $D_{-j,0}^{p}f$; $p = 1,2,3$; $j = 1,2,3\ldots$.

Assuming the number of coefficients in the wavelet filter is small compared with the number of samples N in each dimension of the image sequence, the computational complexity of this stage of the algorithm is found by determining the total number of values computed by the spatial decomposition process. To this end, first note that the number of samples at each spatial decomposition level are half the number at the next higher level. Thus, if the spatial dimensions of a frame in the originally sampled signal are $N \times N$, then the dimensions of a frame at the next lower spatial decomposition level

are $N/2 \times N/2$. Furthermore, since four signals are produced by the spatial decomposition process (1 approximation and 3 detail), the total number of values computed by the first spatial decomposition is $N^2/4 + N^2/4 + N^2/4 + N^2/4 = N^2$. Continuing the process, the next spatial decomposition produces a total of $N^2/4$ values per frame, and so on. Letting the number of spatial decompositions go to infinity then gives an upper bound on the number of spatial values computed *per frame* of $4N^2/3$. Finally, assuming their are N frames in the image sequence, the total number of *spatial* values computed in the spatial decomposition stage of the algorithm is then $N \cdot (4N^2/3) = 4N^3/3$.

In the next stage of the algorithm, the first level spatial detail signals $D_{-1,0}^{p}f$; $p = 1,2,3$ are decomposed in time by convolving flipped versions of the temporal filters $\tilde{h}$ and $\tilde{g}$ across all frames at each spatial location and decimating the temporal dimension by a factor of two (Fig. 10). The temporal decomposition algorithm is then applied in a cascade fashion to each of the temporal approximation signals to yield a set of temporal detail signals, $D_{-1,k}^{p}f$; $p = 1,2,3$; $k = 1,2,3,\ldots$, for each spatial detail signal in the first spatial decomposition level. This process is then repeated for each spatial detail signal $D_{-j,0}^{1}f$, $D_{-j,0}^{2}f$, and $D_{-j,0}^{3}f$; $j = 2,3,4,\ldots$ at each stage of the spatial decomposition process. The temporal detail signals
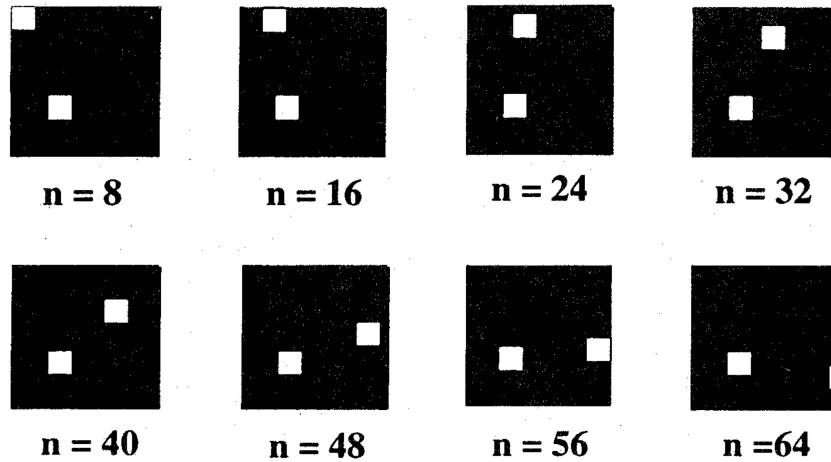
Fig. 11. Several frames of an animated scene consisting of a stationary rectangle and a moving rectangle of equal size and intensity. Each frame contains 64×64 pixels.

produced by this process are represented by the unshaded planes in Fig. 9.

In order to determine the computational complexity of the temporal decomposition stage of the algorithm, note that the upper bound on the number of temporal values computed over all temporal decomposition levels at one spatial location is $2N$. Consequently, given that the number of spatial coefficients produced in the spatial decomposition process is bounded by $4N^2/3$, the total number of coefficients computed in the temporal stage of the algorithm is then $2N \cdot (4N^2/3) = 8N^3/3$. Finally, adding the upper bounds on the spatial and temporal decomposition processes yields an upper bound of $8N^3/3 + 4N^3/3 = 4N^3$ for the total number of values computed in the spatio-temporal decomposition process. Therefore, the computational complexity of the decoupled wavelet subband transform is $\mathcal{O}(N^3)$. The next section provides several results obtained by applying the decoupled fast wavelet transform to various synthetic and real IR image sequences.

VII. RESULTS

The decoupled fast wavelet transform used in these examples was written in C and implemented on a SUN SPARCstation 2. The discrete convolutions in Fig. 10 were carried out with a three-dimensional shift-and-multiply routine. Border problems, which are a common problem in convolution schemes, were reduced by making the borders of each test image sequence symmetric about the spatial and temporal axes.

The first synthetic image sequence consists of a simple animated scene containing a stationary and a moving rectangle of equal size and intensity as shown in Fig. 11. The moving rectangle starts in the upper left corner of the scene and moves to the lower right corner in a parabolic fashion, while the

stationary rectangle remains fixed in the lower left hand corner. Assuming the discretely sampled image sequence represents the approximation coefficients at the resolution level $j = 0$, Fig. 12 shows several frames containing the detail coefficients d^2_{-1}, d^3_{-1}, d^6_{-1}, and d^7_{-1} in equation (10), all of which were produced using a truncated cubic spline (23 coefficients) in space and a Daubechies order 4 wavelet in time [7, 17]. A symmetric cubic spline was selected for the spatial decomposition stage because of its linear phase properties, while the compactly supported Daubechies wavelet was chosen for computational efficiency. Recall that under a homogeneous 3D wavelet multiresolution analysis, the more computationally expensive, truncated infinite-duration impulse response (IIR) cubic spline filter must be used in both space and time.

It is instructive to compare the results in Fig. 12 with those obtained by applying a 2D quad-tree wavelet subband transform to a simple rectangle as demonstrated by Mallat (Fig. 13). As described earlier, under the 2D multiresolution analysis, a rectangle is decomposed into a set of *spatially* oriented frequency channels that capture vertical, horizontal and diagonal features of the image. Fig. 12 shows that under a 3D multiresolution analysis, an image sequence is decomposed into orthogonal *spatio-temporally* oriented frequency channels which now provide the ability to extract these same spatial details for either stationary or moving rectangles.

The next image sequence demonstrates the ability of the unconventional wavelet multiresolution analysis to extract details in an image sequence at different resolutions in time for a fixed spatial resolution. Here, the decoupled fast wavelet transform is applied to an image sequence containing two identical rectangles traveling horizontally with different speeds as shown in Fig. 14.

If the intensity distribution $i(x, y)$ of a single rectangle does not change as it moves across the

n=1 n=12 n=20 n=32

(a)

(b)

(c)

(d)
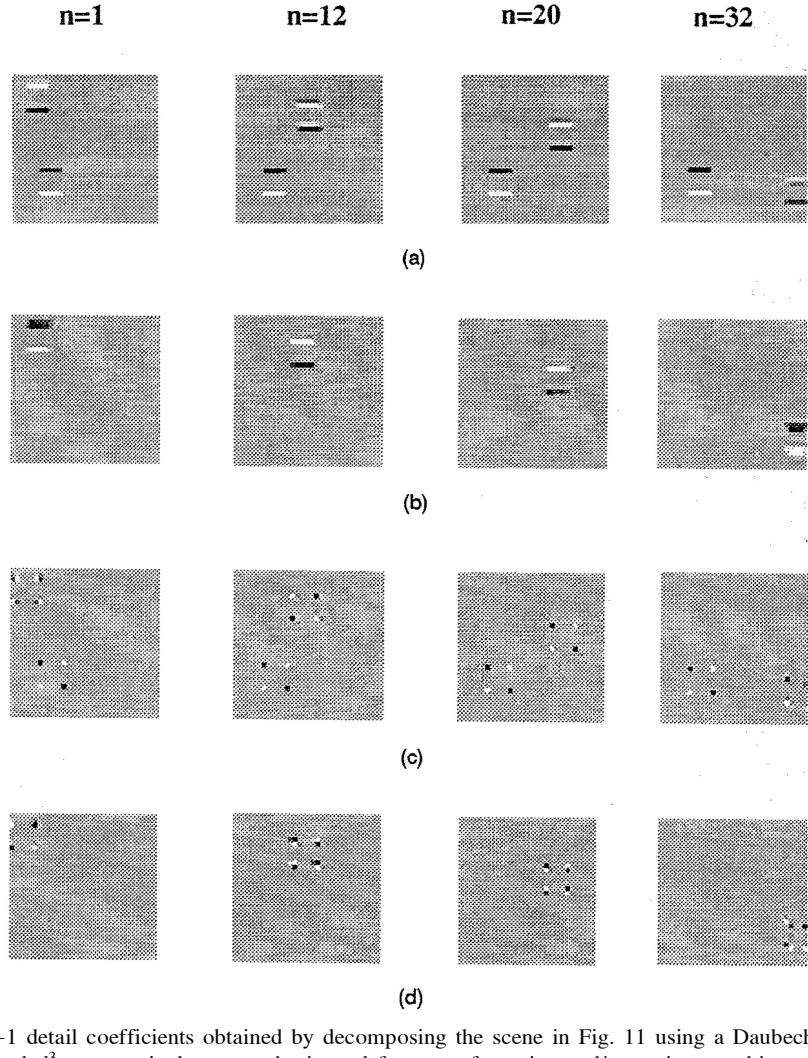
Fig. 12. Four level $j = -1$ detail coefficients obtained by decomposing the scene in Fig. 11 using a Daubechies 4 QMF pair in space and time. In (a) and (b) d^2_{-1} and d^3_{-1}, respectively extract horizontal features of moving and/or stationary objects. In (c) and (d) d^6_{-1} and d^7_{-1}, respectively extract diagonal features of stationary and/or moving objects.

Original Image

Approximation and Detail
Coefficients for 3 Decomposition
Levels
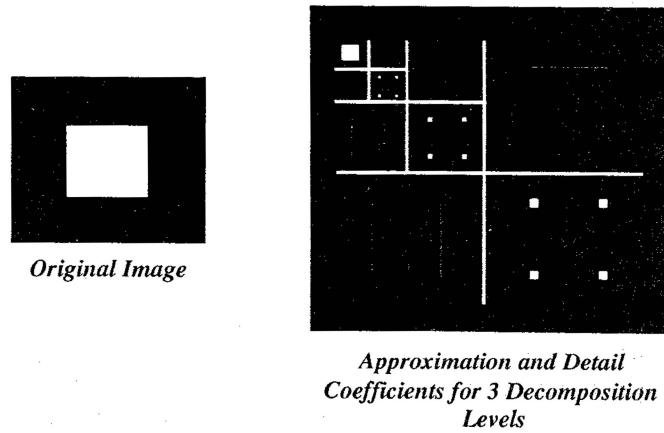
Fig. 13. The magnitude of the coefficients generated by applying Mallat's 2D multiresolution analysis to a simple rectangle [17]. Note that the algorithm extracts horizontal, vertical and diagonal features at each of the three different spatial resolutions. The square in the upper left hand corner is the final approximation signal produced by the decomposition process.

n = 16   n = 32   n = 48   n = 64
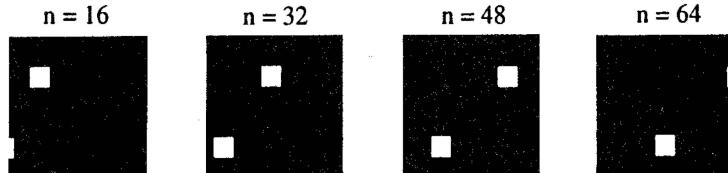
Fig. 14. Several frames of 64×64 synthetic, grayscale imagery containing two equally sized rectangles traveling at different speeds. The speed of the upper rectangle is twice that of the lower rectangle. n represents a frame in the image sequence.


(a)   (b)
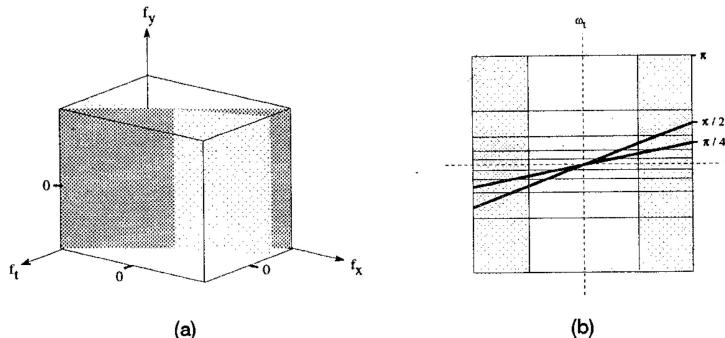
Fig. 15. (a) An idealization of the planes containing the Fourier transforms of the rectangles in Fig. 14. The darker plane corresponds to the faster rectangle. (b) Digital frequency supports of the wavelet filters generated by several decompositions in time for a plane taken through the frequency volume in (a) at the digital spatial frequency $\omega_y = 0$. The dark lines represent the 2D projections of the Fourier transforms of the moving objects. The digital spatial frequency axis ω_y points out of the paper.

image plane (i.e., $i(x - x_0, y - y_0) = i(x, y)$), and if the rectangle moves at a constant velocity, then the 2D Fourier transform of the intensity distribution $I(f_x, f_y)$ is shifted onto a plane in 3D frequency space given by $f_t = -(v_x f_x + v_y f_y)$ where f_x, f_y are spatial frequencies and v_x, v_y are the velocity components of the moving rectangle [12]. In the case of the moving rectangles in Fig. 14, both y velocity components are zero; therefore, their spatio-temporal Fourier transforms lie on the idealized planes shown in Fig. 15(a).

Slicing the frequency volume along the $f_y = 0$ plane yields the two lines shown in Fig. 15(b). The lines contain the spatio-temporal frequency components of the moving rectangles that intersect the spatio-temporal filters in the $f_y = 0$ plane. In order to segment the two horizontally moving objects, Fig. 15(b) suggests choosing the wavelet coefficients at the first spatial decomposition level, and either the second or third temporal decomposition levels. Following this approach, Fig. 16 shows the resulting coefficients for the vertical and diagonal spatial details at the second and third temporal decomposition levels. In this example, a Daubechies 12 QMF pair was used for the temporal decomposition stage of the algorithm because its wide passband and narrow transition region enhances the ability to separate the temporal frequency information in both moving objects. Additionally, spatial frequency resolution is not as critical (i.e., the objects are identical); thus, a Daubechies 4 spatial QMF pair is used to increase the computational efficiency of the spatial decomposition.

The outputs were thresholded to eliminate the small amount of energy captured in overlapping frequency bands of neighboring filters.

The final image sequence, shown in Fig. 17, was chosen to demonstrate the zoom-in and zoom-out capability of the decoupled subband transform. The image sequence contains a large, slow moving tank executing a 180° turn. The imagery is corrupted by background noise, and an exhaust plume is evident behind the tank in frame 100 after it executes the turn. In addition, movements in the camera platform cause the image to jitter slightly from frame to frame.

Since the tank is fairly large and its movement is slow compared with other objects in the scene (notably, the rapidly changing pixels associated with background noise and camera jitter), a large amount of the tank's energy should be contained in the coefficients corresponding to wavelets with longer dilations (i.e., lower resolutions) in space and time. This behavior is clearly evident in Fig. 18 which contains a single frame from each of several different spatial and temporal decomposition levels. By moving horizontally from left to right across the top of the figure (i.e., decreasing the spatial resolution for a fixed temporal resolution), its clear the energy in the spatial wavelet coefficients increases. However, since the temporal resolution is held constant at the highest level, the scintillating background pixels with their correspondingly high temporal frequency energy are still present in the scene. If one now moves

n = 16   n = 32   n = 48

(a)

n = 8   n = 16   n = 24

(b)
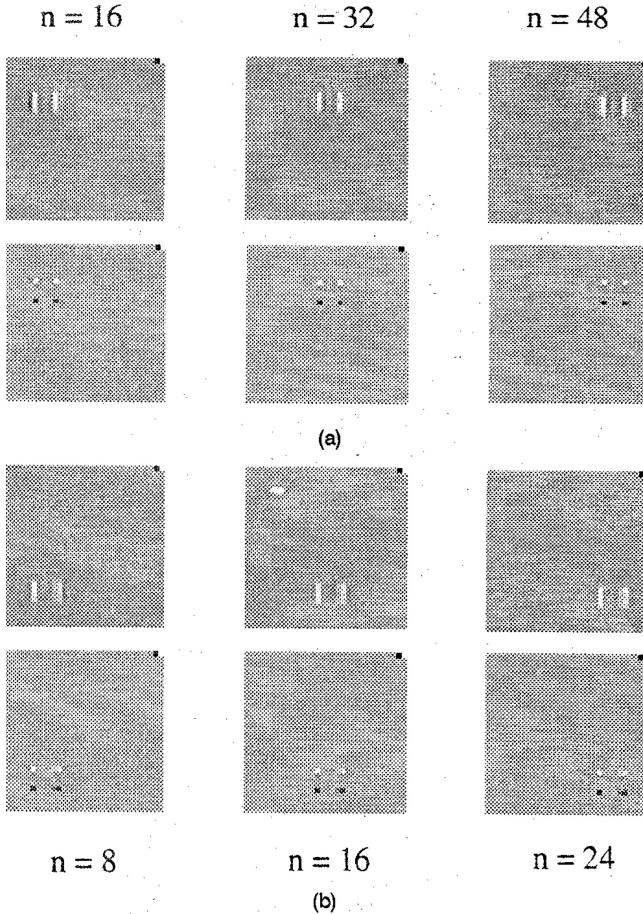
Fig. 16. (a) Segmenting the diagonal and vertical features of the faster object by decomposing the input signal one level in space and one level in time. Slower object completely attenuated by motion-oriented filter bank. The dimensions of the resulting coefficient sequence are $64 \times 64 \times 64$. (b) Segmenting the slower object by decomposing one level in space and two levels in time. The coefficient sequence dimensions are $64 \times 64 \times 32$ (row, column, frame). Here n represents a frame in a coefficient sequence. In this case, the faster object is completely eliminated by the filter bank.

vertically down the right side of the figure, so that the temporal resolution decreases for a fixed spatial resolution, the coefficients associated with the large, slow tank become more and more evident, until, at the spatio-temporal resolution level $j = -3$, $k = -2$, only the blurred tank remains in the image. Thus, a wavelet filter tuned to match the frequency spectrum of large, moderately slow moving objects successfully extracts the tank from the noisy image sequence.

A second test conducted on the tank image sequence was performed to compare the motion segmentation properties of the decoupled multiresolution analysis with a more traditional segmentation technique known as frame differencing. In this technique, pixel values in an image frame at time $t + 1$ are subtracted from the image at time t in order to remove stationary objects from the scene. Two common problems with frame differencing techniques, however, are 1) they require pixel registration between image frames in order to "subtract out" stationary information, and 2) frame-to-frame pixel scintillations caused by noise are not removed by the differencing process. The major advantage of the technique is that it can be implemented in real time at a moderate cost [3].

Fig. 19 compares several unprocessed frames of wavelet detail coefficients to similar poses of the tank produced by a simple frame-differencing operation. Part b shows the wavelet coefficients at the spatio-temporal resolution level $j = 1$, $k = 3$. In each frame, the spatial and temporal detail signals of the vertical, horizontal, and diagonal features are combined to yield a complete outline of the tank. Note that the frame differencing technique shown in part a is quite susceptible to noise sources in the "stationary" background. Conversely, time-averaging properties of the decoupled wavelet decomposition capture the


n=1   n=25   n=50

n=75   n=100   n=125
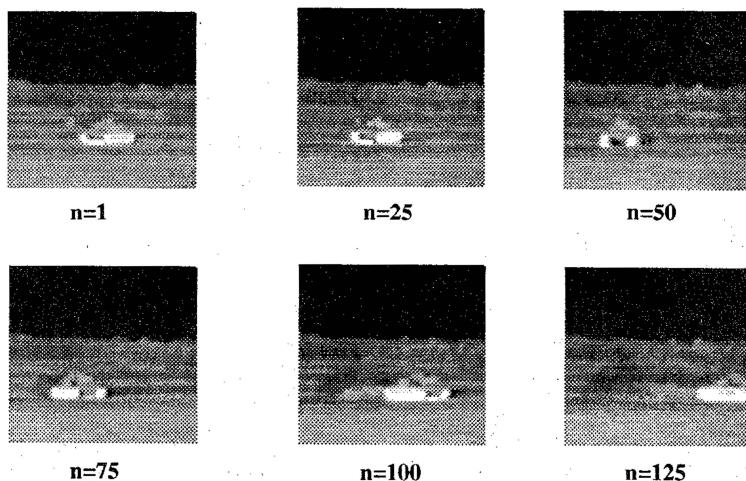
Fig. 17. Several frames of a sequence of IR images in which a large, slow moving tank executes a 180° turn.

Decreasing Spatial Resolution ⟶


Large, Slow Object
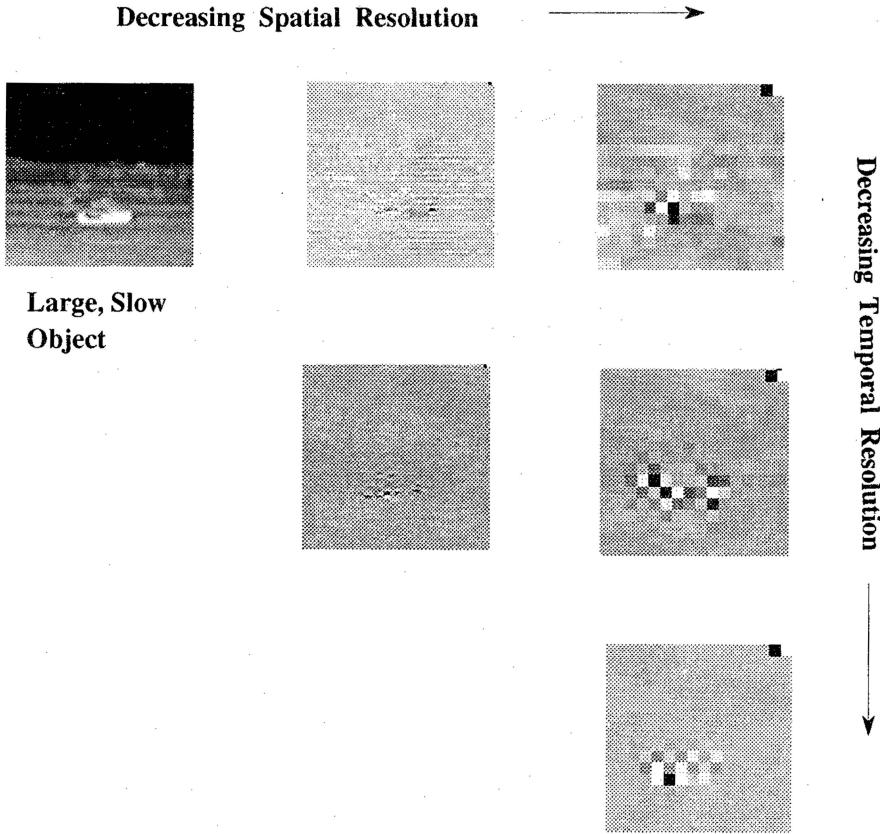
Fig. 18. Single, unprocessed frame of detail coefficients from each of several different motion decompositions in space and time. Moving left to right across the figure increases the spatial dilation of the wavelet which in turn extracts lower spatial frequencies from the sequence. Moving from top to bottom increases the wavelet's temporal dilation, thereby extracting lower temporal frequencies from the sequence. The lower right image corresponds to a spatio-temporal resolution of $j = -3$, $k = -2$.

edges of the tank while virtually eliminating extraneous motion related information in the background.

VIII. DISCUSSION

The spatial decomposition stage of the decoupled subband transform yields three detail sequences at each spatial resolution level that capture horizontal, vertical, and diagonal details in the image sequence for a fixed resolution in time. Many image processing applications, however, require spatial filters that are tunable to a wider range of orientations. Indeed, a considerable amount of evidence indicates biological "image processing" systems perform scale *and* orientation analyses over localized regions in a visual scene [1, 8, 13]. The purpose of this section, therefore, is to briefly present a simple modification of the separable, decoupled subband transform that enhances the orientation selectivity of the spatio-temporal filters.
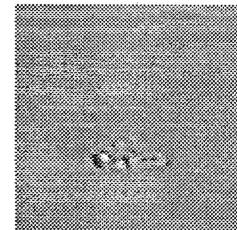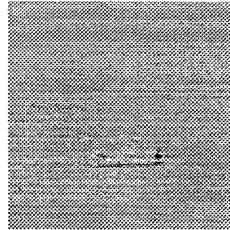
The orientation filtering properties of the separable wavelet filters produced by a 2D quad-tree decomposition structure are limited, as shown previously in Fig. 2, to horizontal, vertical, and diagonal spatial frequency bands at each spatial

resolution level. In order to achieve a higher degree of orientation selectivity, one can construct a nonseparable QMF pair on a hexogonal sampling lattice which, when used in a four-band perfect reconstruction filter bank, produces the orientation specific filters shown in Fig. 20 [23]. These filters can also be extended to three-dimensions to yield garnet shaped filters tuned to different scales, spatial orientations and, to a limited degree, direction of motion [23].
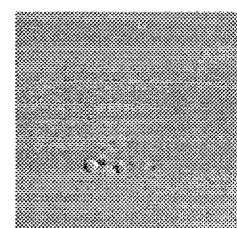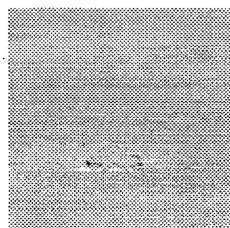
Although the hexogonal nonseparable 3D filters provide a higher degree of orientation selectivity than the separable filters constructed under the decoupled subband transform, they still have several limitations for spatio-temporal signal representation and analysis. First, the impulse responses of the analysis filters are not orthonormal; therefore, one loses the economy of representation provided under an orthonormal series expansion of the signal. Second, like the "conventional" 3D subband transform described earlier, the nonseparable, hexogonal 3D filter bank also restricts the analysis of an image sequence to identical resolutions in space and time. And third, from a practical perspective, nonseparable, multidimensional filter banks are generally more

frame
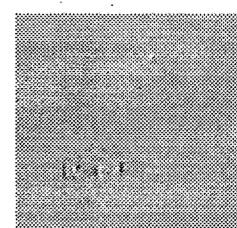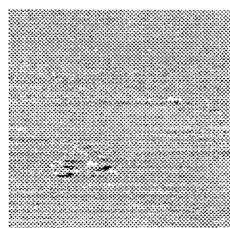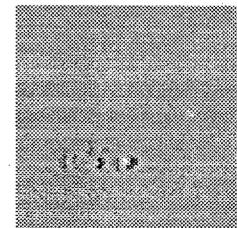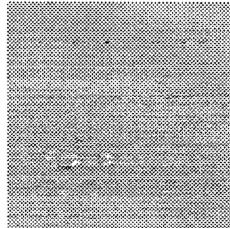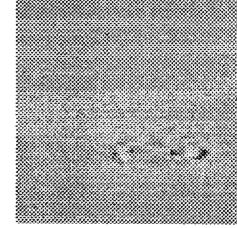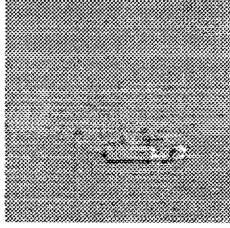
Fig. 19.   (a) Several frames of moving tank image sequence processed with a traditional frame-differencing motion extraction technique. (b) Detail coefficients generated by a motion-oriented wavelet decomposition at the spatial and temporal resolution levels $j = 1$, $k = 3$. The temporal details of the horizontal, vertical and corner spatial details have been combined to form an outline of the moving object.

difficult to construct and implement than their separable counterparts [9].

It is possible, however, to overcome these limitations while retaining the "nice" construction and implementation properties of the separable 3D orthogonal filter bank. Recall that the under the decoupled subband transform the ability to look across multiple resolutions in time for a fixed spatial resolution was achieved by applying Daubechies Lemma to the temporal components of the 3D

separable basis functions of the spatial detail space $\mathbf{W}_{j,k}$. Similarly, it is also possible to apply Daubechies' Lemma to the *spatial* components of these basis functions before decomposing the resulting spatial detail signals in time. Furthermore, the Lemma ensures that the basis functions for these new detail spaces remain orthogonal over all resolutions in space and time. As shown by the idealized filters in Fig. 21, recursively applying Daubechies Lemma in space and time produces an orthogonal filter bank with
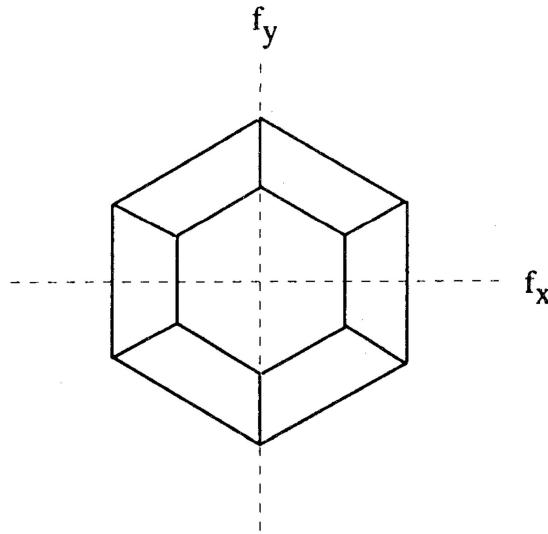
Fig. 20. Idealized supporting regions in 2D frequency space of a "hexogonal" orientation selective filter for a single spatial resolution level [23].

the subband transform used earlier to generate the decoupled filter bank in Fig. 9. Using the new strategy, an image sequence is first decomposed spatially into the horizontal, vertical and diagonal detail sequence described earlier. The horizontal spatial detail sequence is then further decomposed as many times as desired by recursively convolving the spatial QMF pair, h, g, with the columns of each spatial approximation sequence and decimating by a factor of two. Similarly, the spatial filter pair is recursively convolved with the rows of the approximation sequences formed by decomposing the vertical detail sequence. Each of the resulting orientation specific spatial detail sequences are then individually decomposed in time using the temporal decomposition algorithm shown earlier in Fig. 8. The process is then repeated at each successive spatial approximation level.

a higher degree of orientation selectivity (but less spatial resolution!) than the filters constructed with the previous decoupled subband transform.

The subband transform used to construct the filter bank in Fig. 21 is difficult to illustrate in a standard tree diagram. The strategy, however, is quite straightforward assuming one understands

IX. CONCLUSION

This paper presented an unconventional $L_2(\mathbf{R}^3)$ multiresolution wavelet analysis designed for the purpose of representing and analyzing time-sequential imagery. A theoretical framework was first developed that allows for the construction of an $L_2(\mathbf{R}^3)$ multiresolution wavelet analysis from three nonidentical $L_2(\mathbf{R})$ spatial and temporal multiresolution wavelet analyses. This
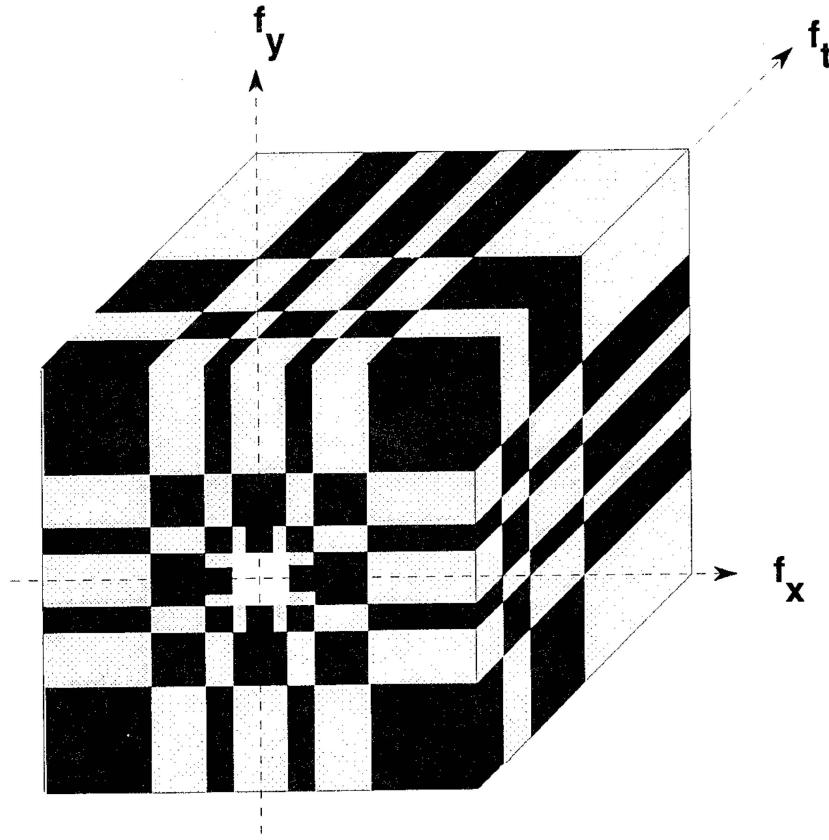

Fig. 21. Idealized supporting regions in the positive temporal half of 3D frequency space of the orthogonal filter bank obtained by applying Daubechies Lemma in space and time.

nonhomogeneous framework provides greater flexibility for tailoring the spatio-temporal frequency characteristics of the three dimensional wavelet filter to match the frequency behavior of the analyzed signal. A decoupled, wavelet subband transform was then described which yields a rich set of *spatio-temporally* oriented frequency channels which, unlike the conventional oct-tree decomposition algorithm, provides independent zoom-in and zoom-out capability in space and time.

The nonhomogeneous decoupled subband transform was applied to a simple image sequence to demonstrate its ability to extract vertical, horizontal, and diagonal features from moving or stationary objects. It was also shown that decoupling the conventional spatial and temporal decomposition processes provides the ability to segment identical objects traveling at different speeds. The algorithm was then applied to a natural IR image sequence to demonstrate its ability to independently zoom-in and zoom-out in space and time to locate objects at different spatial scales in the presence of extraneous motion related phenomena such as camera jitter, background noise, and sensor noise.

Finally, the orientational selectivity of the decoupled subband transform was improved by applying Daubechies' Lemma in space and time. Unlike previously constructed nonseparable perfect reconstruction filter banks, the separable, orientation selective filter bank presented here ensures the corresponding basis functions (i.e., filter impulse responses) remain orthogonal, provides a tool for independently analyzing spatial and temporal details in an image sequence at different orientations and scales, and maintains the efficiency and ease of implementation of a separable multidimensional filter bank.

APPENDIX

This appendix contains key elements of the proofs of Propositions 1–4.

PROOF OF PROPOSITION 1   Let $\Phi_{j;(l,m,n)}(x,y,t) = 2^{3j/2}\phi(2^j x - l)\phi(2^j y - m)\tilde{\phi}(2^j t - n)$. Then

$$\langle \Phi_{j;(l,m,n)}, \Phi_{j;(l',m',n')}\rangle$$
$$= \begin{cases} 1, & \text{if} \quad l = l' \text{ and } m = m' \text{ and } n = n' \\ 0, & \text{otherwise} \end{cases}$$

$$\text{(22)}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product on $L_2(\mathbf{R}^3)$. Equation (22) implies the set of vectors $\{\Phi_{j;(l,m,n)} \mid (l,m,n) \in \mathbf{Z}^3\}$ forms an orthonormal set in $L_2(\mathbf{R}^3)$.

Now, let F be a vector in $\mathbf{V}_j$. By definition of $\mathbf{V}_j$, $F(x,y,t) = f(x)g(y)h(t)$ for some $f \in V_j^x$, $g \in V_j^y$ and $h \in \tilde{V}_j^t$. Expressing f, g, and h in terms of their respective orthonormal bases and rearranging terms

yields

$$F = \sum_l 2^j \langle f, \phi_{j;l}\rangle \phi_{j;l} \sum_m 2^j \langle g, \phi_{j;m}\rangle \phi_{j;m} \sum_n 2^j \langle h, \tilde{\phi}_{j;n}\rangle \tilde{\phi}_{j;n}$$

$$= \sum_l \sum_m \sum_n 2^{3j/2}\langle fgh, \phi_{j;l}\phi_{j;m}\tilde{\phi}_{j;n}\rangle \phi_{j;l}\phi_{j;m}\tilde{\phi}_{j;n}$$

$$= \sum_l \sum_m \sum_n 2^{3j/2}\langle F, \Phi_{j;(l,m,n)}\rangle \Phi_{j;(l,m,n)} \qquad \text{(23)}$$

where $\phi_{j;q} = \phi(2^j \cdot -q)$ and $\tilde{\phi}_{j;q} = \tilde{\phi}(2^j \cdot -q)$. Equation (23) shows F can be expressed as a Fourier series expansion of the orthonormal set $\{\Phi_{j;(l,m,n)} \mid (l,m,n) \in \mathbf{Z}^3\}$. Thus, by the Fourier Series Theorem [19], $\{\Phi_{j;(l,m,n)} \mid (l,m,n) \in \mathbf{Z}^3\}$ forms an orthonormal basis for $\mathbf{V}_j$.

PROOF OF PROPOSITION 2   Recall that a multiresolution analysis of $L_2(\mathbf{R}^3)$ consists of a chain of closed, linear approximation spaces V_j and a scaling function ϕ which satisfy the following properties [5].

1)

$$\cdots V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots. \qquad \text{(24)}$$

2)

$$\overline{\bigcup_{j\in\mathbf{Z}} V_j} = L_2(\mathbf{R}); \qquad \bigcap_{j\in\mathbf{Z}} V_j = \{0\}. \qquad \text{(25)}$$

3)

$$f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}; \qquad j \in \mathbf{Z}$$
$$f(x) \in V_j \Rightarrow f\left(x + \frac{n}{2^j}\right) \in V_j; \qquad n \in \mathbf{Z}. \qquad \text{(26)}$$

4) The set of functions $\{2^{j/2}\phi(2^j x - n) \mid n \in \mathbf{Z}\}$ forms an orthonormal basis for the approximation space V_j.

Proposition 1 ensures Property 4 is met. The intersection condition of Property 2 and Property 3 follow straightforwardly from the fact that the 3D approximation space $\mathbf{V}_j$ is separable. Now show that the denseness condition in Property 2 holds.

Let

$$\mathbf{M} = \overline{\bigcup_{j\in\mathbf{Z}} \mathbf{V}_j} \qquad \text{(27)}$$

and assume $\mathbf{M}$ is not equal to $L_2(\mathbf{R}^3)$. $\mathbf{M}$ is therefore a proper subspace of $L_2(\mathbf{R}^3)$ and, by Hahn-Banach [19], there exists a linear functional ℓ on $L_2(\mathbf{R}^3)$ such that $\ell(M) = 0 \; \forall M \in \mathbf{M}$ and $\ell(G) \neq 0$ for some $G \in L_2(\mathbf{R}^3) - \mathbf{M}$. Then, by the Riesz Representation Theorem [24], there exists a unique $H \in L_2(\mathbf{R}^3)$ such that

$$\ell(F) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(x,y,t)H(x,y,t)\,dx\,dy\,dt$$

$$\text{(28)}$$

$\forall F \in L_2(\mathbf{R}^3)$. Furthermore, if ℓ does not equal the zero functional, then $H \neq 0$. Additionally, $l(M) = 0 \ \forall M \in \mathbf{M}$ implies $H \perp M$. Consequently, the orthogonal projection of H onto $V_j \in M$, P_jH, must equal zero. Now, since $H \in L_2(\mathbf{R}^3)$, there exists a compactly supported C^∞ function, H_0, such that $\|H_0 - H\| < \epsilon$. And, by the Orthogonal Projection Theorem, $\|P_jH_0\| = \|P_j(H_0 - H)\| \leq \|H_0 - H\| < \epsilon$. Thus, by Parseval's Identity,

$$\|P_jH_0\|^2 = 2^{3j/2} \sum_l \sum_m \sum_n |\langle \phi(2^jx - l)\phi(2^jy - m)$$
$$\times \tilde{\phi}(2^jt - n), H_0(x,y,t)\rangle|^2$$
$$\leq \epsilon^2. \tag{29}$$

Using standard mathematical manipulations, it can be shown that

$$\sum_l \sum_m \sum_n |\langle \phi(2^jx - l)\phi(2^jy - m)\tilde{\phi}(2^jt - n), H_0(x,y,t)\rangle|^2 \tag{30}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\hat{H}_0(\xi, \eta, \tau)|^2$$
$$\times |\hat{\phi}(2^{-j}\xi)\hat{\phi}(2^{-j}\eta)\hat{\tilde{\phi}}(2^{-j}\tau)|^2 \, d\xi \, d\eta \, d\tau + R_j \tag{31}$$

where $\hat{H}_0$ denotes the Fourier Transform of H_0 and

$$|R_j| \leq \sum_{l \neq 0} \sum_{m \neq 0} \sum_{n \neq 0} |\hat{H}_0(\xi + 2\pi2^jl, \eta + 2\pi2^jm, \tau + 2\pi2^jn)|$$
$$\times |\hat{f}(\xi, \eta, \tau)|. \tag{32}$$

Now consider the sequence of functions

$$H_j(\xi, \eta, \tau) = \sum_{l \neq 0} \sum_{m \neq 0} \sum_{n \neq 0}$$
$$\times |\hat{H}_0(\times + 2\pi2^jl, \eta + 2\pi2^jm, \tau + 2\pi2^jn)|. \tag{33}$$

Since H_0 is a compact C^∞ function, $\hat{H}_0$ is uniformly bounded and $H_j \to 0$ as $j \to \infty$. Additionally, $\hat{H}_0 \in L_1(\mathbf{R}^3)$ implies $R_j \to 0$ as $j \to \infty$ [5]. Moreover, $\hat{\phi}$ and $\hat{\tilde{\phi}}$ are continuous and uniformly bounded and $\hat{\phi}(0) = \hat{\tilde{\phi}}(0) = 1$. Hence, Lebesgue's Dominated Convergence Theorem can be applied in conjunction with (29) to obtain

$$\lim_{j \to \infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\hat{H}_0(\xi, \eta, \tau)|^2$$
$$\times |\hat{\phi}(2^{-j}\xi)\hat{\phi}(2^{-j}\eta)\hat{\tilde{\phi}}(2^{-j}\tau)|^2 \, d\xi \, d\eta, d\tau$$
$$= \|H_0\|^2 \leq \epsilon^2. \tag{34}$$

Finally, $\|H_0\| \leq \epsilon$ and $\|H_0 - H\| < \epsilon$ implies $\|H\| \leq 2\epsilon$. But ϵ arbitrarily small implies $H = 0$, which contradicts the opening assumption. Thus, $\mathbf{M}$ is dense in $L_2(\mathbf{R}^3)$.

PROOF OF PROPOSITION 3   Let $\mathbf{V}_j$, $j \in \mathbf{Z}$, be a multiresolution approximation of $L_2(\mathbf{R}^3)$ formed by the tensor product

$$\mathbf{V}_j = V_j^x \otimes V_j^y \otimes \tilde{V}_j^t$$
$$= \overline{\text{Span}\{F(x,y,t) = f(x)g(y)h(t) \mid f \in V_j^x, \ g \in V_j^y \ \text{and} \ h \in \tilde{V}_j^t\}} \tag{35}$$

where V_j^x, V_j^y and $\tilde{V}_j^t$ are multiresolution approximations of $L_2(\mathbf{R})$. Let W_j^x, W_j^y, and $\tilde{W}_j^t$ be the orthogonal complements of the closed, linear spaces $V_j^x \subset V_{j+1}^x$, $V_j^y \subset V_{j+1}^y$ and $\tilde{V}_j^t \subset \tilde{V}_{j+1}^t$. Then

$$\mathbf{V}_{j+1} = V_{j+1}^x \otimes V_{j+1}^y \otimes V_{j+1}^t$$
$$= (W_j^x \oplus V_j^x) \otimes (W_j^y \oplus V_j^y) \otimes (W_j^t \oplus V_j^t). \tag{36}$$

The right-hand side of (36) can be rewritten as follows

$$\text{RHS} = [W_j^x \otimes W_j^y \otimes W_j^t] \oplus [W_j^x \otimes W_j^y \otimes V_j^t]$$
$$\oplus [W_j^x \otimes V_j^y \otimes W_j^t] \oplus [W_j^x \otimes V_j^y \otimes V_j^t]$$
$$\oplus [V_j^x \otimes W_j^y \otimes W_j^t] \oplus [V_j^x \otimes W_j^y \otimes V_j^t]$$
$$\oplus [V_j^x \otimes V_j^y \otimes W_j^t] \oplus [V_j^x \otimes V_j^y \otimes V_j^t]. \tag{37}$$

Since $\mathbf{V}_j = V_j^x \otimes V_j^y \otimes V_j^t$, the orthogonal complement, $\mathbf{W}_j$, of $\mathbf{V}_j$ in $\mathbf{V}_jp$ can be expressed as

$$\mathbf{W}_j = \mathbf{V}_jp - \mathbf{V}_j$$
$$= [W_j^x \otimes W_j^y \otimes W_j^t] \oplus [W_j^x \otimes W_j^y \otimes V_j^t]$$
$$\oplus [W_j^x \otimes V_j^y \otimes W_j^t] \oplus [W_j^x \otimes V_j^y \otimes V_j^t]$$
$$\oplus [V_j^x \otimes W_j^y \otimes W_j^t] \oplus [V_j^x \otimes W_j^y \otimes V_j^t]$$
$$\oplus [V_j^x \otimes V_j^y \otimes W_j^t]. \tag{38}$$

The sets of functions $\{2^{j/2}\phi(2^jx - l) \mid l \in \mathbf{Z}\}$, $\{2^{j/2}\phi(2^jy - m) \mid m \in \mathbf{Z}\}$, and $\{2^{j/2}\tilde{\phi}(2^jt - n) \mid n \in \mathbf{Z}\}$, form orthonormal bases, respectively, for the $L_2(\mathbf{R})$ approximation spaces V_j^x, V_j^y, and $\tilde{V}_j^t$. Additionally, the sets of functions $\{2^{j/2}\psi(2^jx - l) \mid l \in \mathbf{Z}\}$, $\{2^{j/2}\psi(2^jy - m) \mid m \in \mathbf{Z}\}$, and $\{2^{j/2}\tilde{\psi}(2^jt - n) \mid n \in \mathbf{Z}\}$, form orthonormal bases, respectively, for the complementary spaces W_j^x, W_j^y, and $\tilde{W}_j^t$. Thus, the set of functions $\{\Psi_j^p(x - l, y - m, t - n) \mid (l,m,n) \in \mathbf{Z}^3; p = 1,2,\ldots,7\}$ forms an orthonormal bases for $\mathbf{W}_j$. Furthermore, the fact that $L_2(\mathbf{R}^3)$ can be formed by the direct sum decomposition

$$\overline{\bigoplus_{j \in \mathbf{Z}} \mathbf{W}_j} = L_2(\mathbf{R}^3) \tag{39}$$

implies the family of functions $\{\Psi_j^p(x - l, y - m, t - n) \mid j \in \mathbf{Z}; (l,m,n) \in \mathbf{Z}^3; p = 1,2,\ldots,7\}$ constitutes an orthonormal basis for $L_2(\mathbf{R}^3)$.

PROOF OF PROPOSITION 4   Consider the following Lemma which describes a special case of Coifman and Meyer's wavelet packet theory as proved by I. Daubechies [6, 7].

LEMMA 1 *Let f be any function such that the $f(t-n)$, $n \in \mathbf{Z}$, are orthonormal. Define the functions*

$$F^1(t) = \sum_n h_n f(t-n)$$

$$F^2(t) = \sum_n g_n f(t-n).$$

(40)

Then $\{F^1(t-2m), F^2(t-2m) \mid m \in \mathbf{Z}\}$ forms an orthonormal basis for $\overline{\mathrm{Span}\{f(t-n) \mid n \in \mathbf{Z}\}}$.

Since the functions $2^{k/2}\tilde{\phi}(2^k t - n)$, $n \in \mathbf{Z}$, are orthonormal, Lemma 1 implies $\{F_k^1(t-2m), F_k^2(t-2m) \mid m \in \mathbf{Z}\}$ forms an orthonormal basis for $\overline{\mathrm{Span}\{\tilde{\phi}(2^k t - n) \mid n \in \mathbf{Z}\}}$ where

$$F_k^1(t) = \sum_n h_n \tilde{\phi}(2^k t - n)$$

$$F_k^2(t) = \sum_n g_n \tilde{\phi}(2^k t - n).$$

(41)

Now, let $\Psi_{j,k}^p(x-l_0, y-m_0, t) = \Psi_j^p(x-l_0, y-m_0) \cdot 2^{k/2}\tilde{\phi}(2^k t)$ where the integer pair $(l_0, m_0) \in \mathbf{Z}^2$ is chosen arbitrarily. Next, define the functions $\Psi_{j,k}^{p1}$ and $\Psi_{j,k}^{p2}$ as follows:

$$\Psi_{j,k}^{p1}(x-l_0, y-m_0, t) = \Psi_j^p(x-l_0, y-m_0)F_k^1(t)$$

$$\Psi_{j,k}^{p2}(x-l_0, y-m_0, t) = \Psi_j^p(x-l_0, y-m_0)F_k^2(t).$$

(42)

Then, the set of functions $\{\Psi_{j,k}^{p1}(x-l_0, y-m_0, t-2n), \Psi_{j,k}^{p2}(x-l_0, y-m_0, t-2n) \mid n \in \mathbf{Z}\}$ forms an orthonormal basis for $\overline{\mathrm{Span}\{\Psi_{j,k}^p(x-l_0, y-m_0, t-n) \mid n \in \mathbf{Z}\}}$. But (l_0, m_0) chosen arbitrarily implies $\{\Psi_{j,k}^{p1}(x-l, y-m, t-2n), \Psi_{j,k}^{p2}(x-l, y-m, t-2n) \mid (l,m,n) \in \mathbf{Z}^3\}$ form an orthonormal basis for $\overline{\mathrm{Span}\{\Psi_{j,k}^p(x-l, y-m, t-n) \mid (l,m,n) \in \mathbf{Z}^3\}}$.

REFERENCES
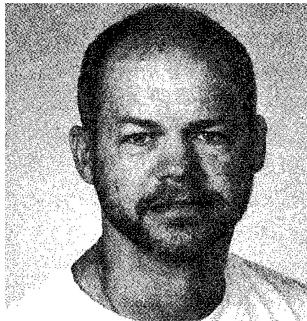
[1] Anderson, C., and Van Essen, D. (1991)
     Processing of visual information in primate brains.
     NASA Tech Brief, **15** (Mar. 1991).

[2] Burns, T. J. (1993)
     A non-homogeneous wavelet multiresolution analysis and its application to the analysis of motion.
     Ph.D. dissertation, School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, Dec. 1993.

[3] Burt, P. J. (1988)
     Smart sensing within a pyramid vision machine.
     Proceedings of the IEEE, **76** (Aug. 1988), 1006–1015.

[4] Chen, T., and Vaidyanathan, P. P. (1993)
     Recent developments in multidimensional multirate systems.
     IEEE Transactions on Circuits and Systems for Video Technology, **3** (Apr. 1993), 116–137.

[5] Chui, C. K. (1992)
     Wavelets: A Tutorial in Theory and Application.
     San Diego: Academic Press, 1992.

[6] Coifman, R. R., and Meyer, Y. (1992)
     Size properties of wavelet-packets.
     In Ruskai, et al., *Wavelets and Their Applications.*
     San Diego: Academic Press, 1992.

[7] Daubechies, I. (1992)
     Ten Lectures on Wavelets.
     Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[8] De Valois, R. L., and Albrecht, D. G. (1985)
     Spatial frequency selectivity of cells in macaque visual cortex.
     Vision Research, **22** (1985), 545–559.

[9] Dudgeon, C. E., and Mersereau, R. M. (1984)
     Multidimensional Digital Signal Processing.
     Englewood Cliffs, NJ: Prentice-Hall, 1984.

[10] Gafni, H., and Zeevi, Y. Y. (1979)
     A model for processing of movement in the visual system.
     Biological Cybernetics, **32** (1979), 165–173.

[11] Goodman, J. (1968)
     Introduction to Fourier Optics.
     New York: McGraw-Hill, 1968.

[12] Heeger, D. (1988)
     Optical flow using spatiotemporal filters.
     International Journal of Computer Vision, (1988), 279–302.

[13] Hubel, D. H., and Wiesel, T. N. (1962)
     Receptive fields, binocular interaction and functional architecture in the cat's visual cortex.
     Journal of Physiology, (1962), 106–154.

[14] Le Chevalier, F., Bobillot, G., and Fugier-Garrel, C. (1978)
     Radar target and aspect angle identification.
     In *Proceedings of the IEEE 1978 International Conference on Pattern Recognition*, (1978), 398–400.

[15] Libby, E. W. (1993)
     Application of sequence comparison methods to multisensor data fusion and target recognition.
     Master's thesis, School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH, July 1993.

[16] Mallat, S. G. (1989)
     Multiresolution approximation and wavelets.
     Transactions on American Mathematical Society, (Sept. 1989), 69–88.

[17] Mallat, S. G. (1989)
     A theory for multi-frequency signal decomposition.
     IEEE Transactions on Pattern Analysis and Machine Intelligence, **11** (July 1989), 674–693.

[18] Meyer, Y. (1986)
     Principe d'incertitude, bases hilbertiennes et albebres d'operateurs.
     Seminaire Bourbaki, **662** (Oct. 1986), 209–223.

[19] Naylor, A. W., and Sell, G. R. (1982)
     Linear Operator Theory in Engineering and Science.
     New York: Springer-Verlag, 1982.

[20] Rioul, O., and Vetterli, M. (1991)
     Wavelets and signal processing.
     IEEE Signal Processing Magazine, (Oct. 1991), 14–38.

[21] Rosenfeld, A. (Ed.) (1984)
     Multiresolution Image Processing and Analysis.
     New York: Springer-Verlag, 1984.

[22] Seibert, M., and Waxman, A. M. (1992)
     Adaptive 3D object recognition from multiple views.
     IEEE Transactions on Pattern Analysis and Machine Intelligence, **14**, 2 (1992), 107–124.

[23] Simoncelli, E. P., and Adelson, E. H. (1990)
     Non-separable extensions of quadrature mirror filters to multiple dimensions.
     Proceedings of the IEEE, **78**, 4 (1990), 652–663.

[24] Taylor, A., and Lay, D. (1982)
      Introduction to Functional Analysis.
      New York: Wiley, 1982.
[25] Vaidyanathan, P. P. (1993)
      Multirate Systems and Filter Banks.
      Englewood Cliffs, NJ: Prentice-Hall, 1993.
[26] Vetterli, M. (1984)
      Multidimensional subband coding: Some theory and
      algorithms.
      Signal Processing, **6** (Feb. 1984), 97–112.

[27] Vetterli, M., Kovacevic, J., and Ugall, D. J. (1990)
      Perfect reconstruction filter banks for hdtv representation
      and coding.
      Signal Processing: Image Communication, **2** (Oct. 1990),
      349–363.
[28] Vetterli, M., et al. (1992)
      Multiresolution coding techniques for digital television. A
      review.
      Multidimensional Systems and Signal Processing, **3** (May
      1992), 161–187.

Thomas J. Burns received the Bachelor of Arts degree in philosophy in 1980
and the Bachelor of Science degree in electrical engineering in 1986, both
from The Ohio State University, Columbus. During 1989–1993 he earned the
Master of Science and Ph.D. degrees in electrical engineering, with emphasis
in electro-optics and signal processing, both from the Air Force Institute of
Technology, Wright-Patterson Air Force Base, OH.
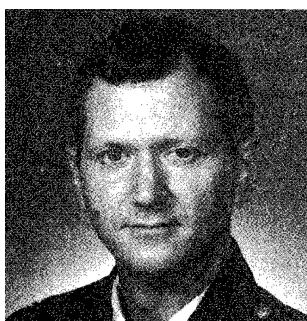
   Major Burns is currently assigned to the Wright Laboratory Automatic Target
Recognition Branch. His interests include model-based vision, spatio-temporal
signal processing, and motion analysis.

Steven K. Rogers is a Professor in the Department of Electrical and Computer
Engineering at the Air Force Institute of Technology, Wright-Patterson Air Force
Base, Dayton, OH.

   He is presently conducting an extensive research program on optical
information processing and neural networks. The research program addresses the
problems inherent in making smart weapons.

   Dr. Rogers has published over one hundred papers in the areas of neural
networks, pattern recognition and optical information processing, and a textbook
titled *Introduction to Biological and Artificial Neural Networks for Pattern
Recognition*. He is an SPIE Fellow.

Dennis W. Ruck received his M.S.E.E and Ph.D. degrees from the Air Force
Institute of Technology, Wright-Patterson Air Force Base, OH, in 1987 and 1990,
respectively, and his B.S.E.E. from Northwestern University, Evanston, IL, in 1983
all in electrical engineering.

   He has been on the faculty at the Air Force Institute of Technology since 1990
as an Assistant Professor. His prior assignment was at the Aeronautical Systems
Division in the visual systems for flight simulators branch. His research interests
include pattern recognition, speech and image processing and recognition, and
neural networks.

   Dr. Ruck has published several articles in the areas of neural networks and
target recognition.

Mark E. Oxley received the B.S. degree in mathematics in 1978 from Cumberland
College, Williamsburg, KY, the M.S. degree in applied mathematics in 1980 from
Purdue University, West Lafayette, IN, and the Ph.D. degree in mathematics in
1987 from North Carolina State University, Raleigh, NC.

   Since 1987, he has been with the Graduate School of Engineering at the Air
Force Institute of Technology, Wright-Patterson Air Force Base, OH, where is
an Associate Professor of Mathematics in the Department of Mathematics and
Statistics. His current research interests include neural networks, wavelet analysis,
functional analysis, and nonlinear partial differential equations.

   Dr. Oxley is a member of Pi Mu Epsilon, the American Mathematical Society
(AMS), the Society for Industrial and Applied Mathematics (SIAM), the American
Geophysical Union (AGU), and SPIE the International Society for Optical
Engineering.