Combining the Discrete Wavelet Transform and Mixed-Domain Filtering

Michael S. Lazar, *Member, IEEE,* and Leonard T. Bruton, *Fellow, IEEE*

***Abstract*—A novel filtering method is proposed that combines the discrete orthogonal wavelet transform (DWT) with the mixed-domain (mixed-D) filtering method. The method uses the DWT to pre- and postprocess those dimensions of the signal that are transformed to the discrete-frequency domain by mixed-D filtering. Using the DWT in this manner provides a controlled mechanism to partition the spectrum of the input signal into subband signals, which then may be selectively filtered during the linear difference equation (LDE) step of the mixed-D algorithm. It is shown that, when the DWT is computed using filters with ideal high- and lowpass frequency responses, the LDE filters used in the mixed-D filtering stage are unchanged by the introduction of the DWT (although the frequency tuple associated with each LDE filter is altered). This indicates that the mixed-D filtering scheme can be easily used in subband coding systems. Results are given for the filtering of a three-dimensional (3-D) linear trajectory signal, representing a common application in video processing.**

I. Introduction

SUBBAND filtering, which decomposes an m-dimensional (m-D) signal into a set of signals corresponding to different frequency bands, has received considerable attention recently [12], [13], [16]. This type of decomposition permits independent processing of each subband based upon the spectral characteristics of the input signal and is advantageous for many practical m-D signals in which most of the spectral energy is contained in the low-frequency bands. Subband filtering is also related to multiresolution decomposition [5] and the discrete wavelet transform (DWT) [1]–[3]. The DWT, which can be used to efficiently perform multiresolution analysis, can be computed by iteratively applying a subband decomposition to an input signal. The output of the DWT consists of a set of signals at various resolutions and scales, corresponding to a self-similar tiling in the m-D frequency space.

Typically, m-D filtering operations are performed by using either spatial domain techniques, employing linear difference equations (LDE's), or by using frequency domain techniques. A novel approach, known as mixed-domain (mixed-D) filtering, *combines these techniques*. This is achieved by taking a p-dimensional discrete Fourier transform (DFT) ($p < m$) of the signal, and then applying the LDE filtering to the transformed p-tuples along the remaining $m - p$ dimensions [10]. It has been shown that mixed-D filtering lends itself to real-time filtering applications, is computationally more efficient than the DFT method, and the resulting mixed-D m-D filter is, in many cases, easier to design than the corresponding m-D LDE filter (particularly with respect to the stability of the resulting filter) [10]. Mixed-D filtering can also be generalized to include various signal transforms and signal processing techniques [11].

In this contribution, we present a novel m-D filtering method that combines multiresolution decomposition, via the wavelet transform, and mixed-D filtering. Specifically, for an m-D signal, a p-dimensional ($p < m$) DWT is applied to the signal, the resulting signals are mixed-D filtered, and then a p-dimensional inverse DWT is used to create the output signal. It is shown, by example, that the frequency selectivity of the DWT may be used to reduce the overall number of computations by filtering only those DWT subsequences that contain significant portions of the signal energy. Given the close tie between the m-D DWT and subband coding systems [4], this work also illustrates how the mixed-D filtering method can be incorporated within recently proposed video coding schemes using the wavelet transform [12], [13], [16].

This paper is organized as follows: A brief review of both mixed-D filtering and the wavelet transform is provided in Section II; in Section III, the proposed method is presented; in Section IV, we present a video processing application of the combined approach, and in Section V conclusions and directions for further work are presented.

II. A Review of Mixed-D Filtering and the Wavelet Transform

In this section, the mixed-D filtering method and the DWT are reviewed. More detailed coverage of these topics can be found in [1]–[4] and [7]–[11].

A. Mixed-D Filtering

Consider the m-dimensional discrete-domain signal, $x(\mathbf{n}^{(m)})$, where the spatiotemporal domain index vector is given by $\mathbf{n}^{(m)} = (n_1, n_2, \cdots, n_m)^T, n_i \in Z, Z = \{0, 1, \cdots\}$, and where ^{T} indicates transposition. Two approaches are often used to filter $x(\mathbf{n}^{(m)})$: block processing using discrete-frequency domain transforms such as the DFT and LDE filtering, which leads to continuous-frequency domain spectra.

Manuscript received September 30, 1993; revised November 1, 1995. This work was supported by MICRONET, the (Canadian) Federal Network of Centres of Excellence on Microelectronics, Devices and Systems, the Natural Sciences and Engineering Research Council of Canada, and the University of Calgary through their Silver Anniversary Scholarship Program.The associate editor coordinating the review of this paper and approving it for publication was Prof. Nikolas P. Galatsanos.

The authors are with the Department of Electrical and Computer Engineering, University of Calgary, Calgary, Alberta T2N 1N4 Canada (e-mail: bruton@enel.ucalgary.ca).

Publisher Item Identifier S 1057-7149(96)04544-7.

The first step in DFT filtering is to convert the input signal, $x(\mathbf{n}^{(m)})$, to the discrete-frequency domain by using the m-D DFT: $X(\mathbf{k}^{(m)}) = \mathrm{DFT}[X(\mathbf{n}^{(m)})]$, where $\mathbf{k}^{(m)} = (k_1, k_2, \cdots, k_m)^T$, $k_i \in Z$, is the m-D discrete-frequency domain index [17]. The transformed data, $X(\mathbf{k}^{(m)})$, are then multiplied by a weighting function, $H(\mathbf{k}^{(m)})$, corresponding to the desired filtering operation, to yield $Y(\mathbf{k}^{(m)}) = X(\mathbf{k}^{(m)})H(\mathbf{k}^{(m)})$. The output is computed by taking the inverse DFT of $Y(\mathbf{k}^{(m)})$, $y(\mathbf{n}^{(m)}) = \mathrm{DFT}^{-1}[Y(\mathbf{k}^{(m)})]$.

In the LDE filtering approach, the input and output signals, $x(\mathbf{n}^{(m)})$ and $y(\mathbf{n}^{(m)})$, are related by the m-D linear difference equation

$$y(\mathbf{n}^{(m)}) = \sum_{\mathbf{i}^{(m)} \in R_a} a(\mathbf{i}^{(m)}) x(\mathbf{n}^{(m)} - \mathbf{i}^{(m)}) - \sum_{\mathbf{i}^{(m)} \in R_b, \mathbf{i}^{(m)} \neq \mathbf{0}^{(m)}} b(\mathbf{i}^{(m)}) y(\mathbf{n}^{(m)} - \mathbf{i}^{(m)}) \quad (1)$$

where R_a and R_b denote the finite regions of support for the input and output m-D difference equation coefficients, $a(\mathbf{i}^{(m)})$ and $b(\mathbf{i}^{(m)})$, respectively, and $\mathbf{0}^{(m)}$ represents an m-tuple with all indexes equal to zero. The frequency response of (1) is determined by the coefficients $a(\mathbf{i}^{(m)})$, $b(\mathbf{i}^{(m)})$, and its domain is continuous and in $\Re^{(m)}$ ($\Re$ denotes the set of real numbers). *Selecting the LDE coefficients to yield stable and useful transfer functions, however, can be a very difficult design problem.*

In mixed-D filtering, the DFT and LDE methods are combined so that LDE filtering takes place on DFT-transformed data [10]. Specifically, a partial p-dimensional DFT, $p < m$, of $x(\mathbf{n}^{(m)})$ is computed over the first p-dimensions of $x(\mathbf{n}^{(m)})$. The partial DFT is given by [10] the following:

$$X(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)}) = \mathrm{DFT}^p[x(\mathbf{n}^{(m)})] = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \cdots \sum_{n_p=0}^{N_p-1} x(\mathbf{n}^{(m)}) \exp[-j2\pi(\mathbf{k}^{(p)})^T \mathbf{L}\mathbf{n}^{(p)}] \quad (2)$$

where $\mathbf{k}^{(p)} = (k_1, k_2, \ldots, k_p)^T, \mathbf{n}^{(m-p)} = (n_{p+1}, n_{p+2}, \cdots, n_m)^T, \mathbf{n}^{(p)} = (n_1, n_2, \cdots, n_p)$, L is a two-dimensional (2-D) diagonal matrix, $L = \mathrm{diag}(N_1, N_2, \cdots, N_p)$, and N_i is the size of $x(\mathbf{n}^{(m)})$ along the ith dimension. The mixed domain index, $(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)})$, indicates that the first p dimensions are in the frequency domain and the remaining $m - p$ dimensions are in the spatiotemporal domain.

The resulting (complex) p-tuples of $X(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)})$ are then filtered using a *distinct* $m - p$ dimensional LDE *for each value of* $\mathbf{k}^{(p)}$. The LDE filtered p-tuples are given by

$$Y(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)}) = X(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)}) * h(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)}) \quad (3)$$

where $*$ represents $(m - p)$-dimensional linear convolution over the variables $\mathbf{n}^{(m-p)}$, and $h(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)})$ is the impulse response of the LDE for the frequency p-tuple $\mathbf{k}^{(p)}$. The output signal, $y(\mathbf{n}^{(m)})$, is obtained by computing the inverse p-dimensional DFT on the LDE filtered p-tuples, $y(\mathbf{n}^{(m)}) = \mathrm{IDFT}^p[Y(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)})]$, where

$$\mathrm{IDFT}^p[Y(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)})] = \frac{1}{|\mathbf{L}|} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \cdots \sum_{n_p=0}^{N_p-1} Y(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)}) \times \exp[j2\pi(\mathbf{k}^{(p)})^T \mathbf{L}^{-1} \mathbf{n}^{(p)}]. \quad (4)$$

It can be shown that the frequency response of the mixed-D filtering method is determined by the $m - p$ dimensional LDE filters $h(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)})$ [10]. Although there may be, in general, a large number of LDE filters (one for each $\mathbf{k}^{(p)}$), their design is an $m - p$ dimensional problem, as opposed to the m-dimensional design problem that is required for the corresponding m-D LDE in (1). However, when the number of LDE filters is very large, the memory required to store the LDE coefficients may make mixed-D filtering prohibitive. Given *a priori* knowledge of the magnitude spectrum of the input signal, the arithmetic complexity of the mixed-D filtering method may be significantly reduced by neglecting many of the p-dimensional LDE filtering operations.

B. The Wavelet Transform

Continuous-time wavelets are basis functions that result from the dilations and translations of a "mother" wavelet, $\psi(t)$ [1]–[3]. For discrete-time (and therefore resolution-limited) one-dimensional (1-D) signals, a tree-structured lossless perfect reconstruction (PR) filter bank can be used to realize an orthogonal discrete wavelet transform (DWT) [3]. Perfect reconstruction filter banks have been developed for subband coding applications and are well studied in the signal processing literature [4]. *Throughout this contribution we will restrict our attention to dyadic orthogonal wavelets* [3], in which case the DWT is computed using a two-band PR filter bank. We denote the impulse response of the lowpass and highpass PR filters by $h_0(n)$ and $h_1(n)$, respectively, and their discrete-time Fourier transforms by $H_0(e^{j\omega})$ and $H_1(e^{(j\omega)})$.

Continuous-domain wavelets can be obtained from the discrete-domain filters $h_0(n)$ and $h_1(n)$ via an infinite cascade of PR filters and subsamplers [5]. It is sometimes desirable that the continuous-domain wavelet associated with $h_0(n)$ and $h_1(n)$ be "regular," i.e., continuously differentiable [3]. For dyadic wavelets, regularity can be achieved by ensuring that $H_0(e^{j\omega})$ contain zeros at $\omega = \pi$, with regularity increasing with the number of these zeros [2].

In order to implement an m-D DWT, the separable extension of the 1-D decomposition is used. It can be shown that for an m-D separable dyadic wavelet transform, there are $2^m - 1$ wavelets [3]. If $B(l, j)$ represents the jth bit in the m-bit binary representation of l, then the corresponding m-D separable PR filters can be written as

$$h_l(\mathbf{n}^{(m)}) = \prod_{j=1}^{m} \begin{cases} h_0(n_j), & \text{if } B(l,j) = 0, \\ h_1(n_j), & \text{if } B(l,j) = 1, \end{cases} \quad l = 0, 1, \cdots, 2^m - 1. \quad (5)$$

Using the PR filters defined by (5), the m-D DWT algorithm can be implemented using separable filtering and subsampling on a rectangularly spaced grid as follows:

$$\begin{aligned} S_{2^{j+1}}x(\mathbf{n}^{(m)}) &= \sum_i h_0(\mathbf{i}^{(m)} - 2\mathbf{n}^{(m)})(S_{2^j}x(\mathbf{i}^{(m)})), \\ S_{2^0}x(\mathbf{n}^{(m)}) &= x(\mathbf{n}^{(m)}), \qquad (6) \\ D^l_{2^{j+1}}x(\mathbf{n}^{(m)}) &= \sum_i h_l(\mathbf{i}^{(m)} - 2\mathbf{n}^{(m)})(S_{2^j}x(\mathbf{i}^{(m)})), \\ & \qquad l = 1, 2, \cdots, 2^m - 1 \qquad (7) \end{aligned}$$

where $S_{2^j}x(\mathbf{n}^{(m)})$ represents the "approximation" of $x(\mathbf{n}^{(m)})$ at the jth resolution, $D^l_{2^j}x(n), l = 1, 2, \cdots, 2^m - 1$, represent the "detail" between the jth and $(j-1)$th resolutions, and where the convolution summations are computed over the m-D region of support for $S_{2^j}x(\mathbf{n}^{(m)})$. The iterative algorithm expressed by (6) and (7) is started from $j = 0$ (the original resolution of $x(\mathbf{n}^{(m)})$ and proceeds to some desired low resolution approximation level, $j = J > 0$ (*increasing* j implies *decreasing* signal resolution in our notation).

The inverse m-D DWT is computed by using the m-D separable extension of the 1-D DWT^{-1} [1] as follows:

$$\begin{aligned} \hat{S}_{2^j}x(\mathbf{n}^{(m)}) = &\sum_i h_0(\mathbf{n}^{(m)} - 2\mathbf{i}^{(m)})\hat{S}_{2^{j+1}}x(\mathbf{i}^{(m)}) \\ &+ \sum_{l=1}^{2^m-1}\sum_i h_l(\mathbf{n}^{(m)} - 2\mathbf{i}^{(m)})D^l_{2^{j+1}}x(\mathbf{i}^{(m)}) \end{aligned} \qquad (8)$$

for $j = J-1, J-2, \cdots, 0$, with the initial condition $\hat{S}_{2^J}x(\mathbf{n}^{(m)}) = S_{2^J}x(\mathbf{n}^{(m)})$, and where the convolution summations are over the regions of support for $S_{2^j}x(\mathbf{n}^{(m)})$ and $D^l_{2^{j+1}}x(\mathbf{n}^{(m)})$. For noncausal filtering, the use of PR filters ensures that $\hat{S}_{2^j}x(\mathbf{n}^{(m)}) = S_{2^j}x(\mathbf{n}^{(m)})$; if m-D causal filters are used, the reconstructed signal, $\hat{S}_{2^j}x(\mathbf{n}^{(m)})$, is an m-D shifted version of $S_{2^j}x(\mathbf{n}^{(m)})$.

III. Combined DWT/Mixed-Domain Filtering

One of the difficulties in using mixed-D filtering is the potentially large number of different required LDE filters. As stated earlier, this number may be reduced by *a priori* knowledge of the frequency domain characteristics of the input signal. For example, if there is little or no signal energy at certain frequencies, then LDE filtering at those frequencies may be skipped. *By preprocessing the input signal with the DWT, we introduce a controlled method of partitioning the frequency spectrum of the input signal prior to mixed-D filtering.* If a particular DWT signal contains little energy, then LDE filtering of that signal can be neglected, resulting in reductions in both the number of LDE's and the overall computation time. Note that, for practical signals, the low-resolution approximation signal will often contain most of the signal energy and, thus, the mixed-D filtering step may be omitted for the difference signals. mixed-D filtering is thus well suited for coding and compression systems that utilize the DWT or subband filtering [12], [13], [16].

In this section, a method is presented whereby the m-D DWT and mixed-D filtering methods are combined. The proposed method, referred to herein as combined DWT/mixed-D filtering, consists of three basic steps: A p-dimensional, $p < m$, DWT of an m-D input signal, $x(\mathbf{n}^{(m)})$, is first computed; mixed-D filtering is then performed on the resulting approximation and difference signals; and finally, the mixed-D filtered approximation and difference signals are combined by performing an inverse p-dimensional DWT. After the proposed method is presented in Section III-B, an equivalent frequency domain description of the algorithm is provided in Section III-C. This frequency domain description is then used in Section III-D to outline some of the considerations involved during LDE filter design in combined DWT/mixed-D filtering. In Section IV, a three-dimensional (3-D) filtering example using the proposed approach is presented, and some computational considerations in its implementation are discussed.

A. Issues Related to Finite-Length Signals

For most practical m-D applications, the input signal, $x(\mathbf{n}^{(m)})$, is unbounded in at most one dimension (typically corresponding to time). In those dimensions that are finitely bounded, we compute the m-D subsample and convolution operations in (6) and (7) using circular convolution with periodically extended versions of the filter coefficients and signals. This ensures that both the PR property of the DWT algorithm and critical sampling rates are maintained [15] (although there may be extraneous energy in the high-frequency bands due to any discontinuities introduced by the assumed periodicity—the same is true for DFT filtering). *Throughout the rest of the paper it is assumed that all filtering operations performed in finitely bounded dimensions are carried out using circular convolution and periodically extended signals.*

B. Proposed Algorithm

The steps in the proposed combined DWT/mixed-D filtering method are illustrated in Fig. 1 and are as follows.

a) Apply a DWT to the first p-dimensions of $x(\mathbf{n}^{(m)})$.
b) Apply a p-dimensional DFT to the approximation and detail signals obtained in (a).
c) LDE filter the DWT/DFT transformed approximation and detail signals along the remaining $m-p$ dimensions.
d) Apply an inverse p-dimensional DFT to the LDE filtered signals from (c).
e) Apply an inverse p-dimensional DWT to the signals from (d) giving $y(\mathbf{n}^{(m)})$.

Step a) of the proposed algorithm can be computed by extending the DWT algorithm into the mixed-domain using

$$\begin{aligned} &S_{2^{j+1}}x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)}) \\ &\quad = \sum_i h_0(\mathbf{i}^{(p)} - 2\mathbf{n}^{(p)})S_{2^j}x(\mathbf{i}^{(p)}, \mathbf{n}^{(m-p)}) \end{aligned}$$

with

$$S_{2^0}x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)}) = x(\mathbf{n}^{(m)}) \qquad (9)$$

and

$$\begin{aligned} &D^l_{2^{j+1}}x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)}) \\ &\quad = \sum_i h_l(\mathbf{i}^{(p)} - 2\mathbf{n}^{(p)})S_{2^j}x(\mathbf{i}^{(p)}, \mathbf{n}^{(m-p)}), \\ &\qquad\qquad l = 1, 2, \cdots, 2^p - 1 \qquad (10) \end{aligned}$$

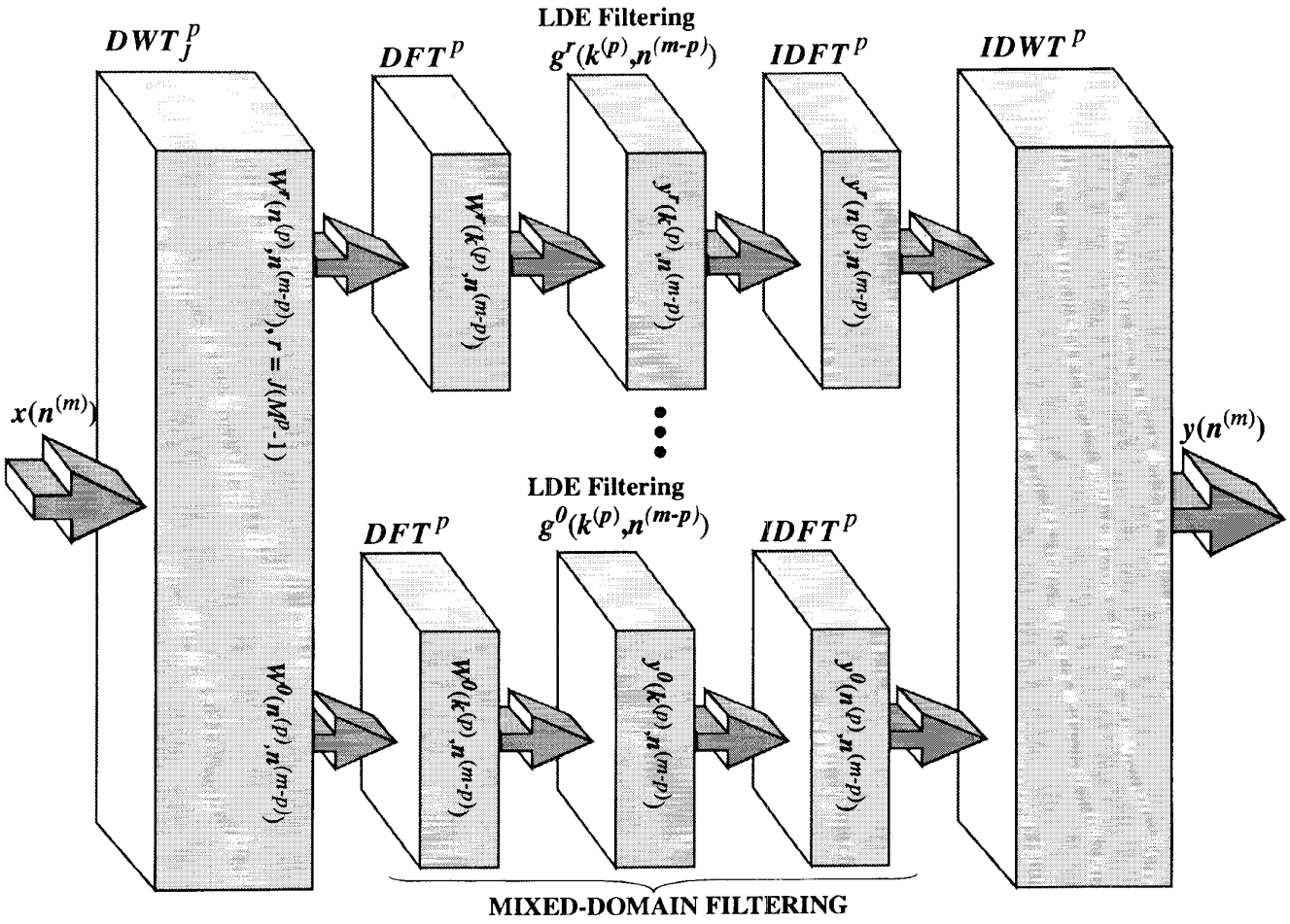

Fig. 1. Combined DWT/mixed-D filtering algorithm.

for $j = 0, 1, \cdots, J-1$, where the (circular) convolution summations are over the regions of support for $S_{2^j}x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)})$ in the first p-dimensions. The index $(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)})$ indicates that the first finitely bounded, periodically extended, p-dimensions are filtered and subsampled while the remaining $m-p$ dimensions are unfiltered. (Therefore, the number of samples in the first p-dimensions can be determined from the operator S_{2^j} (or D_{2^j}); the remaining dimensions retain their original size.)

Mixed-D filtering, corresponding to steps b)–d) above, is applied to the low-resolution approximation signal, $S_{2^J}x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)})$, and the set of difference signals, $\{D^l_{2^j}x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)})\}$, $l = 1, 2, \cdots, 2^p - 1$, $j = 1, 2, \cdots, J$, generated by the DWT. In order to simplify notation, $S_{2^J}x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)})$ and $\{D^l_{2^j}x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)})\}$ are combined into one set, $\{W^r x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)})\}$, defined by

$$W^r x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)}) = \begin{cases} S_{2^J}x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)}) & r = 0, \\ D^l_{2^j}x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)}) & r = 1, 2, \cdots, J(2^p - 1) \end{cases} \quad (11)$$

with difference signal index $l = [(r-1) \bmod (2^p - 1)] + 1$, ($a \bmod b$ is the remainder left after dividing a by b), and resolution index $j = \lceil r/\cdots \rceil (\lceil a/b \rceil$ is the least integer not smaller than a/b).

The first step in mixed-D filtering is to compute the set of filtered subband signals, $\{W^r X(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)})\} = \mathrm{DFT}^p[\{W^r x(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)})\}]$. This step can be accomplished via (2), where the summation bounds are changed to reflect the reduced signal sizes.

For convenience, we denote the set of mixed-D LDE impulse responses used to filter $\{W^r X(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)})\}$ by $g^r(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)})$ where $g^r(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)})$ contains one unique LDE filter for each frequency p-tuple, $\mathbf{k}^{(p)}$). With this notation, the LDE filtering operation of step (c) can be expressed

$$Y^r(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)}) = W^r X(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)}) * g^r(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)}), \quad \forall \mathbf{k}^{(p)},\ r = 0, 1, \cdots, J(2^p - 1) \quad (12)$$

where the (linear) convolution is over the $m-p$ dimensions, $\mathbf{n}^{(m-p)}$.

The final step in mixed-D filtering the DWT signals is to inverse DFT the LDE filtered signals

$$y^r(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)}) = \mathrm{IDFT}^p[Y^r(\mathbf{k}^{(p)}, \mathbf{n}^{(m-p)})]. \quad (13)$$

Equation (13) can be computed via (4) using appropriately sized summation bounds.

Step e) in combined DWT/mixed-D filtering recombines the mixed-D processed signals in the p mixed-D filtered

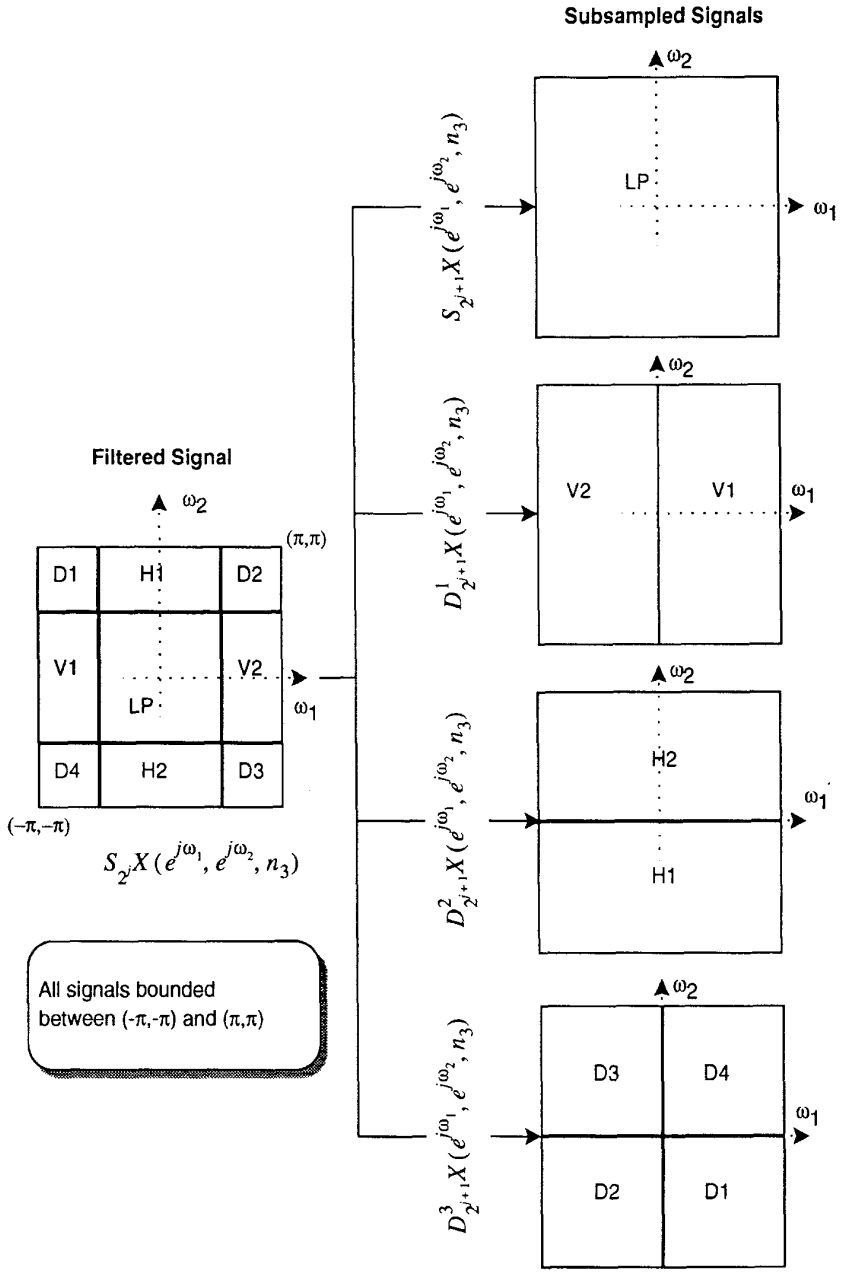

Fig. 2. Redistribution of signal energy in $\omega_1 - \omega_2$ plane during one iteration of the wavelet transform (assuming ideal PR filters).

dimensions by using the inverse p-dimensional DWT, $y(\mathbf{n}^{(m)}) = \mathrm{IDWT}^p[Y^r(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)})]$. The algorithm to compute $\mathrm{IDWT}^p[\cdot]$ can be expressed by

$$\begin{aligned} S_{2^j} y(\mathbf{n}^{(m)}, \mathbf{n}^{(m-p)}) = & \sum_i h_0(\mathbf{n}^{(p)} - 2\mathbf{i}^{(p)}, \mathbf{n}^{(m-p)}) \\ & \times S_{2^{j+1}} y(\mathbf{i}^{(p)}, \mathbf{n}^{(m-p)}) \\ & + \sum_{l=1}^{2^p-1} \sum_i h_l(\mathbf{n}^{(p)} - 2\mathbf{i}^{(p)}, \mathbf{n}^{(m-p)}) \\ & \times y^{j(2^p-1)+l}(\mathbf{i}^{(p)}, \mathbf{n}^{(m-p)}) \end{aligned} \quad (14)$$

for $j = J-1, J-2, \cdots, 0$, with $S_{2^J} y(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)}) = y^0(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)}), y(\mathbf{n}^{(m)}) = S_{2^0} y(\mathbf{n}^{(p)}, \mathbf{n}^{(m-p)})$, and where the (circular) convolution summations are over the regions of support for $y(\mathbf{i}^{(p)}, \mathbf{n}^{(m-p)})$ in the first p-dimensions.

C. Design Issues and the Frequency Domain Description of Combined DWT/Mixed-D Filtering

It is well known that filtering and subsampling operations can be performed in the frequency domain as well as in the spatiotemporal domain. Since, for sufficiently large sample sizes, it is computationally advantageous to perform filtering

Fig. 3. Frame 220 of Traffic video sequence.

operations directly in the m-D DFT domain [17], we investigate how the DWT may be computed in the discrete-frequency domain. In doing so, insight is obtained into the frequency-domain operation of combined DWT/mixed-D filtering. This insight is subsequently useful in considering issues related to LDE filter design.

We denote the Fourier transform of the time-domain filters, $h_i(\mathbf{n}^{(p)}), i = 0,1,\cdots,2^p-1$, by $H_i^{2^j}(\mathbf{k}^{(p)}) = \mathrm{DFT}^p[h_i^{2^j}(\mathbf{n}^{(p)})]$, where the superscript 2^j indicates that the time domain filters are zero padded (before transforming) to contain the same number of samples as the approximation or difference signals at the jth resolution. Using the m-D extension of the DFT subsampling operation [18], the following equations, iterated from $j = 0$ to $j = J-1$, can be used to determine the partial p-dimensional DWT in the discrete-frequency domain

$$\begin{aligned} &S_{2^{j+1}}X(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)}) \\ &\quad = \frac{1}{|\det \mathbf{M}|}\sum_{\mathbf{i}^p=\mathbf{0}^p}^{\mathbf{M}} H_0^{2^j}(\mathbf{k}^{(p)}+\mathbf{M}^{-1}\mathbf{N}(\mathbf{i}^{(p)})^T) \\ &\qquad \times S_{2^j}X(\mathbf{k}^{(p)}+\mathbf{M}^{-1}\mathbf{N}(\mathbf{i}^{(p)})^T,\mathbf{n}^{(m-p)}) \end{aligned} \tag{15}$$

with $S_{2^0}X(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)}) = X(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)})$, and

$$\begin{aligned} &D_{2^{j+1}}^l X(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)}) \\ &\quad = \frac{1}{|\det \mathbf{M}|}\sum_{\mathbf{i}^p=\mathbf{0}^p}^{\mathbf{M}} H_l^{2^j}(\mathbf{k}^{(p)}+\mathbf{M}^{-1}\mathbf{N}(\mathbf{i}^{(p)})^T) \\ &\qquad \times S_{2^j}X(\mathbf{k}^{(p)}+\mathbf{M}^{-1}\mathbf{N}(\mathbf{i}^{(p)})^T,\mathbf{n}^{(m-p)}) \end{aligned} \tag{16}$$

for $l = 1,2,\cdots,2^p-1$. Both $\mathbf{M}$ and $\mathbf{N}$ are two-dimensional (2-D) matrices, $\mathbf{M} = \mathrm{diag}(M_1,\cdots,M_p), \mathbf{N} = \mathrm{diag}(N_1,\cdots,N_p)$, where M_i denotes the decimation rate in the ith dimension and where N_i is the size of $S_{2^j}X(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)})$ in the ith dimension. The m-D summations in (15) and (16) are understood to be over M_i terms in the ith dimension, $i = 1,2,\cdots,p$. It is important to note that the multiplication operations in (15) and (16) are carried out over *each* of the *frequency* p-tuples corresponding to $S_{2^{j+1}}X(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)})$, with the size of $S_{2^{j+1}}X(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)})$ in each frequency dimension given by $k_i = 0,1,\cdots,(N_i/M_i)-1$ for $i = 1,2,\cdots,p$ (we assume that N_i and M_i divide evenly).

Using the m-D extension of the DFT upsampling operation, allows the inverse DWT of the LDE filtered signals, $Y^r(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)})$ [given by (12)] to be expressed in the discrete-frequency domain. If $\hat{H}_l^{2^j}(\mathbf{k}^{(p)})$ denotes the DFT of the "spatiotemporal" reversed version of $h_l^{2^j}(\mathbf{n}^{(p)})$ (defined in (5)), then

$$\begin{aligned} &S_{2^j}Y(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)}) \\ &\quad = \hat{H}_0^{2^j}(\mathbf{k}^{(p)})S_{2^{(j+1)}}Y(\mathbf{k}^{(p)} \bmod (\mathbf{NM}^{-1}),\mathbf{n}^{(m-p)}) \\ &\qquad + \sum_{l=1}^{2^p-1}\hat{H}_l^{2^j}(\mathbf{k}^{(p)}) \\ &\qquad \times Y^{j(2^p-1)+1}(\mathbf{k}^{(p)} \bmod (\mathbf{NM}^{-1}),\mathbf{n}^{(m-p)}) \end{aligned} \tag{17}$$

iterated from $j = J-1$ to $j = 0$ can be used to compute $Y(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)}) = S_{2^0}Y(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)}) = \mathrm{IDWT}^p[Y^r(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)})]$. Here, the multiplication operations in (17) are over *each* of the *frequency* p-tuples corresponding to $S_{2^j}X(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)})$, with the size of each frequency dimension given by $k_i = 0,1,\cdots,N_i-1$. The output signal from the combined DWT/mixed-D filtering is given by returning to the spatiotemporal domain via the inverse DFT, $y(\mathbf{n}^{(m)}) = \mathrm{IDFT}^p[Y(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)})]$.

D. The Design of the Frequency Response in Combined DWT/Mixed-D Filters

It is clear that the frequency response of a combined DWT/mixed-D filter depends not only upon the LDE filters, $g^r(\mathbf{k}^{(p)}\mathbf{n}^{(m-p)})$, but also upon the responses of the PR filters and on the aliasing and frequency shifting caused by subsampling. In this section, it is shown that for PR filters having ideal frequency responses, there is, in fact, no change in the frequency response of the combined DWT/mixed-D filter compared to the corresponding mixed-D filter.

During application of the DWT within combined DWT/mixed-D filtering, each frequency p-tuple $\mathbf{k}^{(p)}$ in the (original or) approximation signal at the jth resolution is distributed, via the subsampling operation, to 2^p-1 ($j \neq J$) or 2^p ($j = J$) other subband signals at the $(j+1)$th resolution. Consider the use of PR filters, $H_l^{2^j}(\mathbf{k}^{(p)})$, having ideal lowpass and highpass frequency responses within the DWT algorithm. For such filters, no aliasing occurs during subsampling and, thus, only the frequency shifting caused by subsampling is of concern. Specifically, each $\mathbf{k}^{(p)}$ is mapped to a single frequency p-tuple within *one* subband signal from the set $\{W^rX(\mathbf{k}^{(p)},\mathbf{n}^{(m-p)})\}$ (corresponding to the filter for which $\mathbf{k}^{(p)}$ resides in the passband).

Note that, for a mixed-D signal, frequency shifting a p-tuple $\mathbf{k}^{(p)}$ *does not effect the spectral energy in the remaining* $m-p$ *dimensions, in which LDE filtering occurs. Thus, for*

ideal PR filters, the same LDE filters can be used in the combined DWT/mixed-D filter as in the corresponding mixed-D filter. The difference is that the LDE filters of the combined DWT/mixed-D filter are associated with different frequency p-tuples than in the mixed-D filter. This indicates that the combined DWT/mixed-D filter can be made to have the same frequency response as the corresponding mixed-D filter (for a discussion of the frequency response of mixed-D filters, the reader may refer to [10]).

In practice, of course, nonideal PR filters are used to compute the DWT. The above discussion suggests, therefore, that PR filters with good passband and stopband responses should be used in combined DWT/Mixed-D filtering to ensure that the LDE filter design step remains unchanged from that of the corresponding mixed-D filter.

Finally, it is important to note that for orthogonal dyadic PR filters (which have been assumed), the filters $h_0(n)$ and $h_1(n)$ cannot have linear phase [3]. Therefore, if the phase of the combined DWT/mixed-D filter is of concern, the phase distortion introduced by the PR filters should be considered during the design of the LDE filters.

IV. Using the Combined DWT/Mixed-D Filtering to Enhance 3-D Linear Trajectory Digital Video Signals

In this section, we apply combined DWT/mixed-D filtering to the practical problem of filtering 3-D linear trajectory (LT) digital video signals. LT signals typically represent objects travelling at a constant velocity along a straight line in the 3-D spatiotemporal domain. The processing and tracking of these signals is becoming increasingly important in many areas such as sonar, computer vision, and video compression [9].

Our model for discrete-domain digital video signals is $x(n_1, n_2, n_3)$, where the first two dimensions are bounded, and the third dimension, corresponding to time, is unbounded; that is

$$x(n_1, n_2, n_3) \mid 0 \le n_1 \le N_1 - 1 < \infty, \quad 0 \le n_2 \le N_2 - 1 < \infty, \ 0 \le n_3 \le \infty. \tag{18}$$

The frame at "time" n_3 is denoted by $x(\cdot, \cdot, n_3)$.

A. Review of the Frequency-Domain Properties of 3-D Linear Trajectory Signals

Consider, for the moment, a 3-D continuous-domain signal, $x_c(t_1, t_2, t_3)$. If $x_c(t_1, t_2, t_3)$ is an LT signal, then, by definition, there exists a direction in 3-D space in which $x_c(t_1, t_2, t_3)$ is constant valued [7]. It can be shown that the magnitude of the 3-D Fourier transform of $x_c(t_1, t_2, t_3)$, $|X_c(j\omega_1, j\omega_2, j\omega_3)|$, is contained within a 3-D plane passing through the origin, with normals given by the 3-D velocity of the LT signal [7], [20], [21] as follows:

$$\alpha_1\omega_1 + \alpha_2\omega_2 + \alpha_3\omega_3 = 0 \tag{19}$$

where α_i is the velocity of the object in the ith dimension, and the vector $(\alpha_1, \alpha_2, \alpha_3)$ defines the normal of the signal plane. It is therefore possible to selectively enhance (or attenuate) an LT signal by using a 3-D filter whose passband (or stopband) is aligned with the signal plane described by (19).

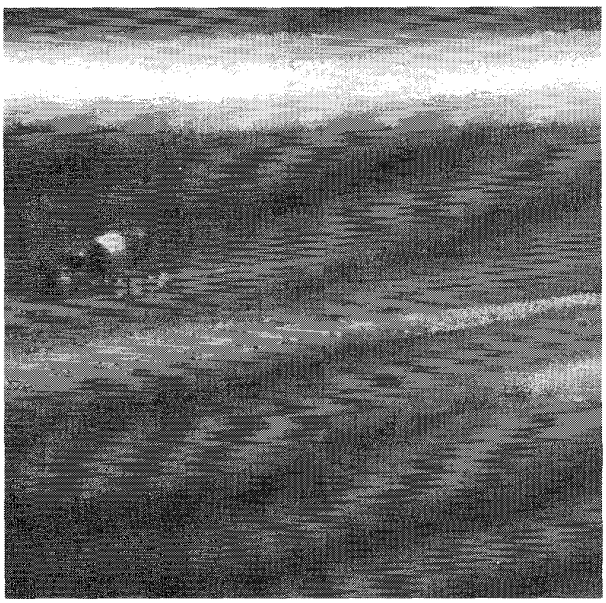

Fig. 4. Frame 220 of mixed-D filtered output.

The discrete-domain signal $x(n_1, n_2, n_3)$ is obtained from $x_c(t_1, t_2, t_3)$ by spatial truncation and sampling. The spectrum, $X(e^{j\omega_1}, e^{j\omega_2}, e^{j\omega_3})$ of the discrete-time Fourier transform (DTFT) of $x(n_1, n_2, n_3)$ is therefore given by the periodic replication of the signal planes with periodicity $(2\pi, 2\pi, 2\pi)$. In other words, $X(e^{j\omega_1}, e^{j\omega_2}, e^{j\omega_3})$ consists of a series of signal planes with normals $(\alpha_1, \alpha_2, \alpha_3)$, such that each plane passes through at least one of the points $(k2\pi, l2\pi, m2\pi), k, l, m \in Z$. If the temporal and spatial sampling rates are sufficiently high, then the spectral energy within each plane will be confined to nonoverlapping volumes of size $2\pi \times 2\pi \times 2\pi$, centered at the points $(k2\pi, l2\pi, m2\pi), k, l, m \in Z$.

There are two aliasing-related effects that arise from sampling LT signals. In the first, inadequate spatiotemporal sampling may cause spectral energy to "leak" out *along* the signal plane and exit the associated volume bounded by the 3-D folding frequencies [20]. This is analogous to undersampling a 1-D continuous-time signal, where the high-frequency energy of the continuous-time signal ends up past the folding frequency, resulting in aliasing. However, the fact that the signal energy is concentrated in a (2-D) plane within the 3-D space provides an extra degree of freedom not available in the 1-D situation. As a result, *depending upon the orientation of the signal plane, it is possible that leakage from one spectral replication may not intersect the signal planes of adjacent replications, and thus the aliasing energy due to this leakage may be recovered.* A necessary (but not sufficient) condition to recover this aliasing energy is that the LT signal plane must not be parallel to the $\omega_1 - \omega_2$ plane (i.e., the object must have nonzero velocity).

The second type of aliasing-related effect arises from the discrete-domain grid used when sampling the LT signal. Specifically, even though the continuous-domain LT signal, $x_c(t_1, t_2, t_3)$, is constant valued in a particular direction in 3-D space (by definition), it is *unlikely* that the rectangularly

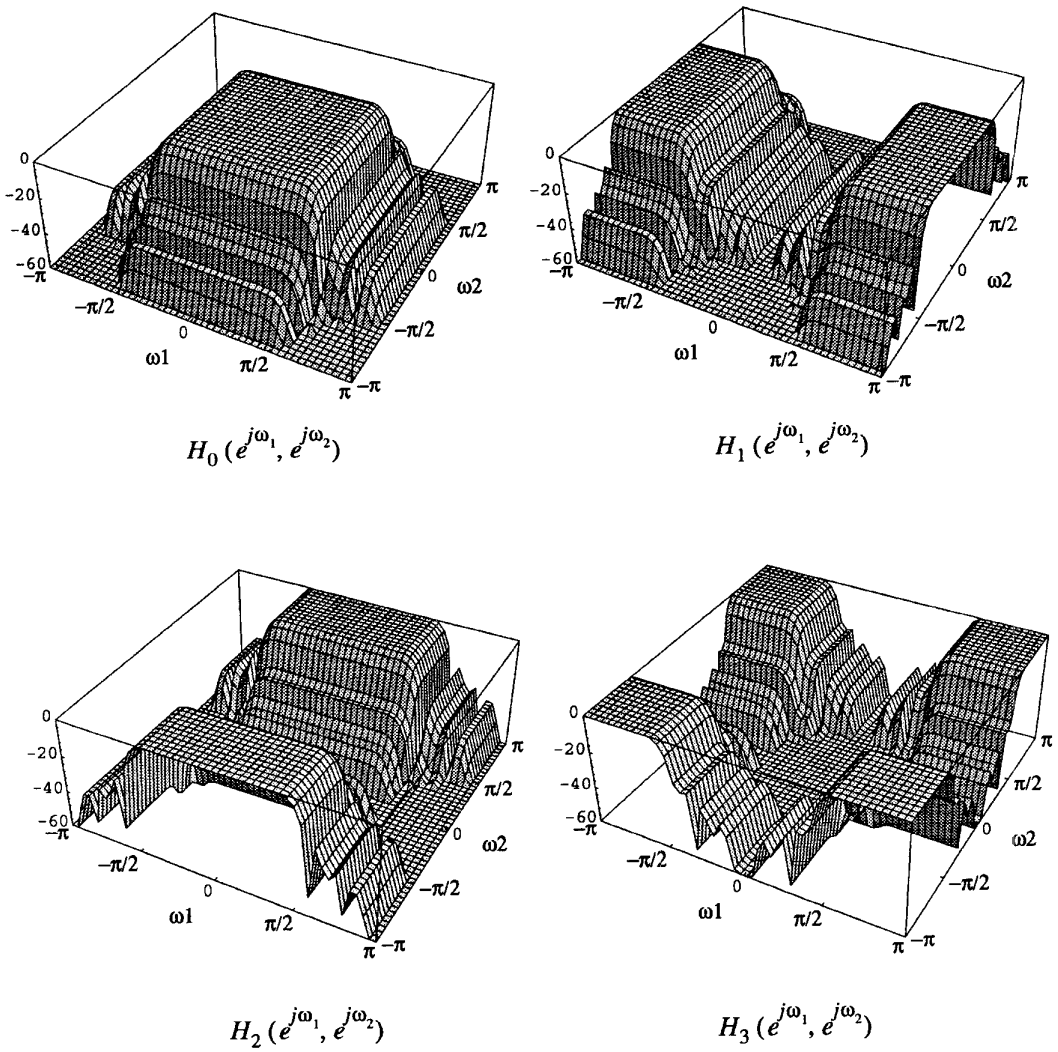

Fig. 5. Magnitude frequency response (in dB) of 2-D PR filters used in LT filtering example.

spaced samples of $x(n_1, n_2, n_3)$ are *exactly* constant valued for those samples representing the trajectory of the LT object. This is due to the fact that it is unlikely, in general, there exists a set of samples that are exactly coincident with the trajectory of the LT object. This effect causes the spectral energy of $x(n_1, n_2, n_3)$ to "smear out" to the neighborhood around the signal plane. It is for this reason that, in practice, wedge shaped 3-D filters are used to filter discrete-domain LT signals [8], [9].

B. Combined DWT/Mixed-D Filtering of 3-D Linear Trajectory Signals: Theory

Consider using the combined DWT/mixed-D filtering method to enhance an LT signal with velocity $v_0 = (c, r)$, where c and r are the horizontal (column) and vertical (row) velocity (in pixels/frame), respectively. After computing the 2-D DWT to J resolution levels, the velocity of the LT signal within *each* subband signal at the jth resolution is (assuming that interpixel distances remain the same at all resolutions) as follows:

$$v_{2^j} = \left(\frac{c}{2^j}, \frac{r}{2^j}\right), \quad 0 \le j \le J. \tag{20}$$

Using similar reasoning as above, the LDE filters used at the jth resolution should be designed to filter objects travelling with velocity given by (20); that is, in the signal plane[1]

$$\frac{c}{2^j}\omega_1 + \frac{r}{2^j}\omega_2 + \omega_3 = 0 \tag{21}$$

with the corresponding frequency relationship

$$\omega_i = \begin{cases} \frac{2\pi k_i}{(N_i/2^j)} & 0 \le k_i \le (N_i/2^j)/2 \\ \frac{2\pi(k_i - (N_i/2^j))}{(N_i/2^j)} & (N_i/2^j)/2 < k_i < (N_i/2^j), \end{cases} \quad i = 1, 2. \tag{22}$$

However, we must be careful when using (20) and (21). It appears initially that, since the velocity of the object is the same in all subband signals, *all* subband signals at a particular resolution could be filtered with the *same* set of LDE filters, thus resulting in fewer distinct LDE filters than in

[1] We assume normalized interframe spacing, $\alpha_3 = 1$.

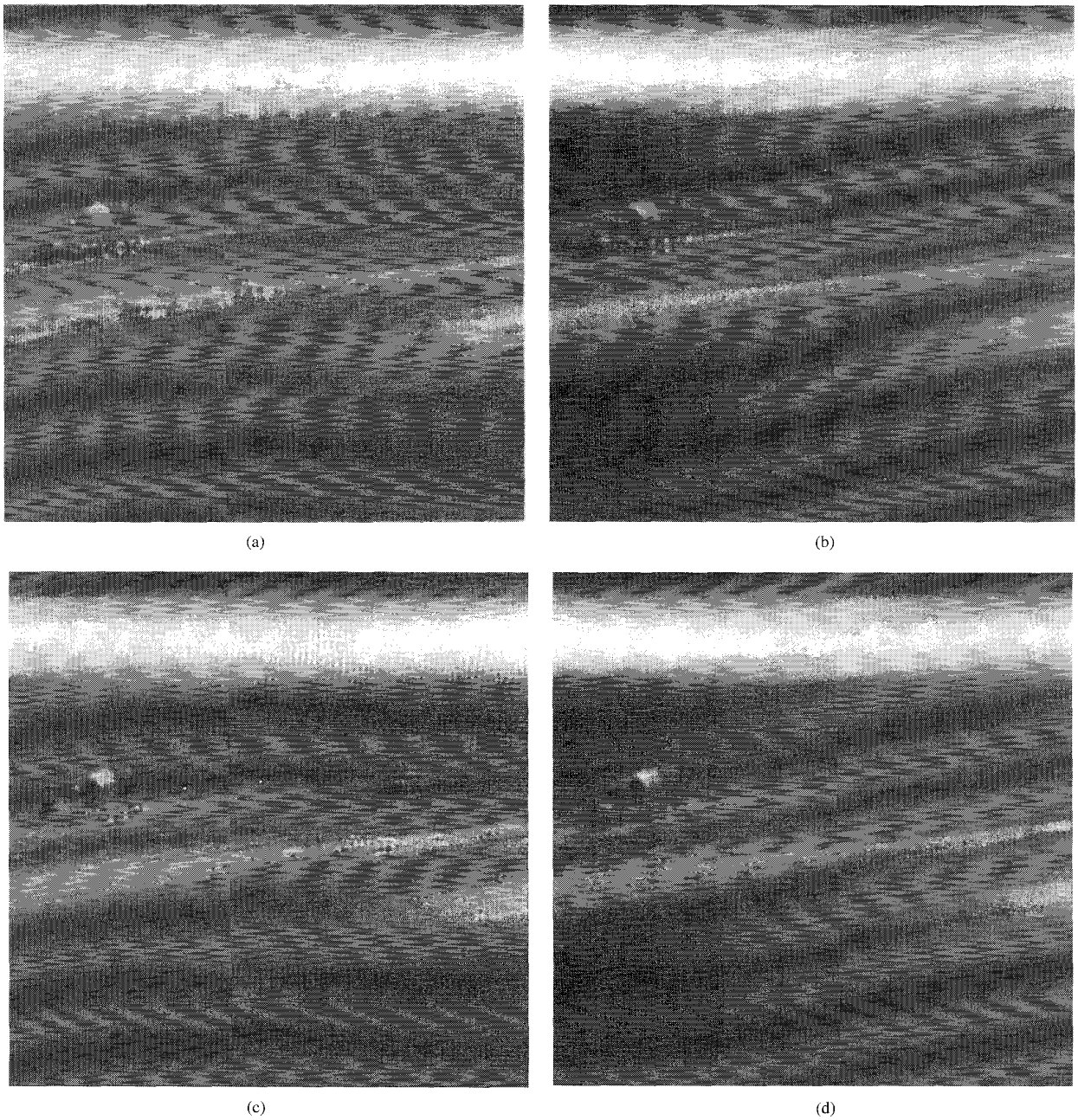

Fig. 6. Frame 220 filtered using the combined DWT/mixed-D filtering method; $p = 2, J = 1$. (a) Mode A-aliasing-compensated filters used to filter all four subbands. (b) Mode B-nonaliasing-compensated filters used to filter all four subbands. (c) Mode C-detail signals unprocessed during mixed-D stage. (d) Mode D-detail signals cleared during mixed-D stage.

the mixed-D method. However, from the discussion regarding LDE filter design, it is known that the *same* number of LDE filters are required in both the combined mixed-D/DWT and mixed-D approaches (assuming ideal PR filters). In order to resolve this apparent contradiction, the effects of spatial filtering and subsampling on a $S_{2^j}X(k_1, k_2, n_3)$ [(15) and (16)] corresponding to an LT signal must be considered.

The filtering and subsampling effects of the 2-D wavelet transform in the continuous-frequency domain are shown in Fig. 2. (The subsampled signals are drawn to the same size as the original to indicate that interpixel distances remain the same). The sampled signals, $S_{2^{j+1}}X(k_1, k_2, n_3)$ and $D_{2^{j+1}}X(k_1, k_2, n_3)$, are obtained from the continuous-domain signals via (22). It is important to note that the distance between frequency domain samples in the ω_1 and ω_2 directions for $S_{2^{j+1}}X(k_1, k_2, n_3)$ and $D_{2^{j+1}}X(k_1, k_2, n_3)$ is double that for $S_{2^j}X(k_1, k_2, n_3)$ (resulting in signals that are one quarter the size).

TABLE I
OPERATIONS OF THE COMBINED DWT/MIXED-D METHOD RELATIVE TO MIXED-D FILTERING

Test Mode	Additional Operations of Proposed Method Relative to Mixed-D Filtering		
	Complex Additions	**Complex Multiplications**	**Complex Storage Values**
Mode A	$12C_{2^1} + 3C_{2^0} - 4C_{2^0} + 64C_{2^1}$ $= 9(N_1 + N_2 + N_1N_2)$	$8C_{2^0} - 5C_{2^0} + 80C_{2^1}$ $= \frac{23}{2}(N_1 + N_2 + N_1N_2)$	$11C_{2^0} - 5C_{2^0} + 80C_{2^1}$ $= 13(N_1 + N_2 + N_1N_2)$
Mode B	$12C_{2^1} + 3C_{2^0} - 4C_{2^0} + 16C_{2^1}$ $= 3(N_1 + N_2 + N_1N_2)$	$8C_{2^0} - 5C_{2^0} + 20C_{2^1}$ $= 4(N_1 + N_2 + N_1N_2)$	$11C_{2^0} - 5C_{2^0} + 20C_{2^1}$ $= \frac{11}{2}(N_1 + N_2 + N_1N_2)$
Mode C	$12C_{2^1} + 3C_{2^0} - 4C_{2^0} + 4C_{2^1}$ $= \frac{3}{2}(N_1 + N_2 + N_1N_2)$	$8C_{2^0} - 5C_{2^0} + 5C_{2^1}$ $= \frac{17}{8}(N_1 + N_2 + N_1N_2)$	$11C_{2^0} - 5C_{2^0} + 5C_{2^1}$ $= \frac{29}{8}(N_1 + N_2 + N_1N_2)$
Mode D	$3C_{2^1} - 4C_{2^0} + 4C_{2^1}$ $= -\frac{9}{8}(N_1 + N_2 + N_1N_2)$	$2C_{2^0} - 5C_{2^0} + 5C_{2^1}$ $= -\frac{7}{8}(N_1 + N_2 + N_1N_2)$	$2C_{2^0} - 5C_{2^0} + 5C_{2^1}$ $= -\frac{7}{8}(N_1 + N_2 + N_1N_2)$

As stated in Section III-D, spatial filtering and subsampling of $S_{2^j}X(k_1, k_2, n_3)$ with adequate PR filters does not change the spectral energy *in the temporal* (n_3) *direction*, but relocates this energy according to how the spatial frequencies are redistributed during subsampling. For an LT signal, the spectral energy of $S_{2^j}X(e^{j\omega_1}, e^{j\omega_2}, n_3)$ in the temporal direction can be thought of as a plane passing through the origin of the $\omega_1 - \omega_2$ plane, where the distance of the plane above (or below) a point in the $\omega_1 - \omega_2$ plane gives the frequency of the temporal energy, and hence indicates the center frequency for the corresponding LDE filter. Fig. 2 can thus be used to illustrate how this plane is "redistributed" throughout the subband signals after subsampling.

For the lowpass region of $S_{2^j}X(e^{j\omega_1}, e^{j\omega_2}, n_3)$ (indicated as LP in Fig. 2), the signal energy is stretched out by a factor of two in both directions. This stretching makes the signal plane appear as if the velocity has been halved in both directions, consistent with both (20) and (21). Note, however, $S_{2^{j+1}}X(k_1, k_2, n_3)$ is obtained from $S_{2^{j+1}}X(e^{j\omega_1}, e^{j\omega_2}, n_3)$ with double the sample spacing used for $S_{2^j}X(k_1, k_2, n_3)$. Thus, there is no actual difference in the corresponding values of ω_3 between the lowpass region of $S_{2^j}X(k_1, k_2, k_3)$ and the signal $S_{2^{j+1}}X(k_1, k_2, n_3)$. Consequently, the same LDE filters are required to filter the lowpass region of $S_{2^{j+1}}X(k_1, k_2, n_3)$ and the signal.

Consider, however, what occurs in the highpass regions of $S_{2^j}X(k_1, k_2, n_3)$. Similar to the lowpass signal, each of the highpass signals is also stretched by a factor of two in both directions, consistent with the observation that the velocity of the LT signal is halved (see (20)). The shifting introduced by the subsampling operation, however, does not move the temporal energy into the plane described by (21). It can be shown, though, *that the frequency shifting of subsampling causes the temporal energy of the detail signals to move into the aliasing planes of* (21) [22]. Similar to the lowpass region described above, the increased sampling distance of $D_{2^{j+1}}X(e^{j\omega_1}, e^{j\omega_2}, n_3)$ ensures that there is no difference in the LDE filters required for the highpass regions of $S_{2^j}X(k_1, k_2, n_3)$ and the signals $D^l_{2^{j+1}}X(k_1, k_2, n_3)$. A more detailed treatment of LT aliasing and its effects on the combined DWT/mixed-D algorithm can be found in [22].

If it is desirable to use the *same* LDE filters when filtering *all* subbands at a particular resolution, aliasing-compensated filters must be used. (Such an approach may be used to reduce resource requirements by, for example, employing a single mixed-D processor on a time-shared basis to filter all subbands at a given resolution.) Note that this approach does not change the total number of LDE filters over the mixed-D method, but it does increase the total number of LDE filtering operations. Depending upon the orientation of the signal plane, using aliasing-compensated filters on the subband signals may further recover some aliased energy from the original signal. Moreover, if the energy of the LT signal is concentrated within the lowpass region of the $\omega_1 - \omega_2$ plane, the total number of LDE filters may be reduced by mixed-D filtering all subbands using LDE filters based on (21).

The total number of filtering operations may be reduced further by LDE filtering only the lowpass approximation signal and leaving intact (or zeroing) the detail signals. When the inverse DWT is performed at the end of mixed-D filtering, the high-frequency information will either be added back to the signal (if left intact), or removed from the signal (if zeroed).

C. The Combined DWT/Mixed-D Filtering of a 3-D Linear Trajectory Signal: A Digital Video Example

To test the combined DWT/mixed-D algorithm, we have obtained a video sequence that is of nine seconds duration, containing moving vehicles [19]. The sequence has been

recorded at the rate of 30 frames per second ($N_3 = 270$) and is of size 320 × 240 pixels. A compressed digital version of the signal has been obtained by first JPEG compressing each frame using JPEG compression parameter $Q = 100$. The signal $x(n_1, n_2, n_3)$ has been obtained by extracting a 220 × 220 section of JPEG decompressed frames, and interpolating (using bilinear interpolation) to frames of size $N_1 = 256 \times N_2 = 256$. The 220th frame, $x(\cdot, \cdot, 220)$ is shown in Fig. 3. The velocity of the cement truck within $x(n_1, n_2, n_3)$ has been estimated to be $v_0 = (-2.5, 0.27)$ (in pixels/frame).

For comparative purposes, $x(n_1, n_2, n_3)$ has been mixed-D filtered using complex coefficient second-order LDE filters (as reported in [8]). Since the input signal is real-valued, the number of unique LDE filters is given by $N_{\text{LDE}} = (N_1N_2)/2 + N_1/2 + N_2/2 = 33\,024$. Frame 220 of the mixed-D filtered series is shown in Fig. 4.

We have processed the sequence $x(n_1, n_2, n_3)$ using combined DWT/mixed-D filtering with $p = 2$, and $J = 1$ (resulting in four subband signals). The 2-D separable DWT has been computed using 20-tap (i.e., FIR) dyadic 1-D PR filters, which have been designed using the Type 2 method, described in [6]. The 1-D lowpass filter $H_0(e^{j\omega})$ contains four zeros at $\omega = \pi$ and thus satisfies the 1-D wavelet regularity requirement [3]. In addition, $H_0(e^{j\omega})$ contains six more zeros on the unit circle that are used to provide improved stopband response. The remaining zeros of $H_0(e^{j\omega})$ have been selected so that the PR property is satisfied. The magnitude responses of the corresponding 2-D separable filters, $H_l(e^{j\omega_1}, e^{j\omega_2})$, $l = 0, 1, 2, 3$, are shown in Fig. 5. (Note that it is possible to use *any* PR filters with good passband and stopband responses to create the subband signals).

The proposed DWT/mixed-D filtering method has been tested using four different filtering modes, herein referred to as Modes A–D. In all test modes, second-order LDE filters, designed using the methods and parameters (except velocities) reported in [8] have been used. The LDE filters have not been adjusted to account for any phase distortion introduced by the 2-D DWT. This is due to the fact the *mixed-D LT filtering algorithm is velocity dependent and position independent, and thus the mixed-D filtering stage is insensitive to any shift of the LT object within the subband signals.*

In test Mode A, all four subband signals have been filtered with the same set of LDE filters. The LDE filters have been designed for the reduced velocity, given by (21), and include LDE filters for the immediate aliasing planes. Thus, the same number of LDE filters are used in Mode A as in mixed-D filtering alone; that is, $N^A_{\text{LDE}} = 33\,024$. Note that Mode A uses approximately four times the number of LDE filtering operations, since the same filtering has been performed in *all* subbands. Frame 220 from test Mode A is shown in Fig. 6(a).

In test Mode B, all subband signals have been also filtered with the same set of LDE filters. In this mode, however, LDE aliasing-compensating filters are not used; therefore, $N^B_{\text{LDE}} = 8256$. Mode B thus requires approximately the same number of LDE filtering operations as in mixed-D filtering alone. Frame 220 from test Mode B is shown in Fig. 6(b).

In the last two modes, the lowpass approximation signal has been LDE filtered without using aliasing compensation, while the mixed-D processing step has been skipped for the highpass detail signals; therefore, $N^C_{\text{LDE}} = N^D_{\text{LDE}} = 8256$. In the test Mode C, the detail signals are passed through unprocessed from the 2-D DWT to the 2-D DWT^{-1} stages, while in test mode D, zero-valued detail signals are used in the 2-D DWT^{-1} step. *Both of these modes thus require approximately one-quarter the number of LDE filtering operations compared to mixed-D filtering alone.* Furthermore, in the test Mode D, the DWT computations involved in creating the detail signals are also not required. Frames 220 from these test modes are shown in Fig. 6(c) and (d).

Comparing the frames shown in Fig. 6 to that shown in Fig. 4 illustrates that the proposed DWT/mixed-D method successfully filters the LT object and yields results that are of comparable subjective quality to those obtained using mixed-D filtering. This is further verified when the sequences are played at video rates. These subjective results are also supported in Fig. 7, where peak signal-to-noise ratio (PSNR) values (relative to the mixed-D filtered sequence) are shown for all four test modes.

The results from test Mode A, which used aliasing-compensated filters, contain some extra-high-frequency detail when compared to the other modes and the mixed-D filtering method. Mode A, however, introduces some high-frequency artifacts that are noticeable as vertical striping.[2] This indicates that the spectrum of the *tested* LT object is indeed well concentrated close to the signal plane associated with the lowpass subband signal. This is further confirmed by the results shown for the other test modes. Specifically, the high PSNR values found when the detail signals have been cleared during the mixed-D stage of processing (Mode D) and when the detail signals have been filtered without aliasing-compensating filters (Mode B) indicate that there is little energy contained in the detail signal subbands. Furthermore, when the detail signals have been passed through the mixed-D stage unprocessed (Mode C), little additional high-frequency information is found in the LT object.

D. Complexity Issues

In this section, the *additional* arithmetic and storage complexity introduced by combined DWT/mixed-D filtering method when compared to the conventional mixed-D filtering method is considered. (For analysis of the number of operations required by mixed-D filtering, the reader may refer to [10].) We restrict our attention to the same dimensions, parameters, and test modes as used in the above example, i.e., $m = 3, p = 2, J = 1$. It is assumed that the input signals are real valued, of even length, and that the 2-D DWT operations are computed in the 2-D discrete-frequency domain, as described in Section III-C.

As noted earlier, there are $C_{2^0} = (N_1N_2 + N_1 + N_2)/2$ unique frequency 2-tuples within $X(k_1, k_2, n_3)$, each requiring a unique LDE filter. For real coefficient PR filters, it is assumed that each subband signal obtained from $X(k_1, k_2, n_3)$

[2] Although not shown here, we have mixed-D filtered the input the signal using aliasing-compensating filters and found the results to also include some additional high-frequency detail and a similar vertical striping artifact.

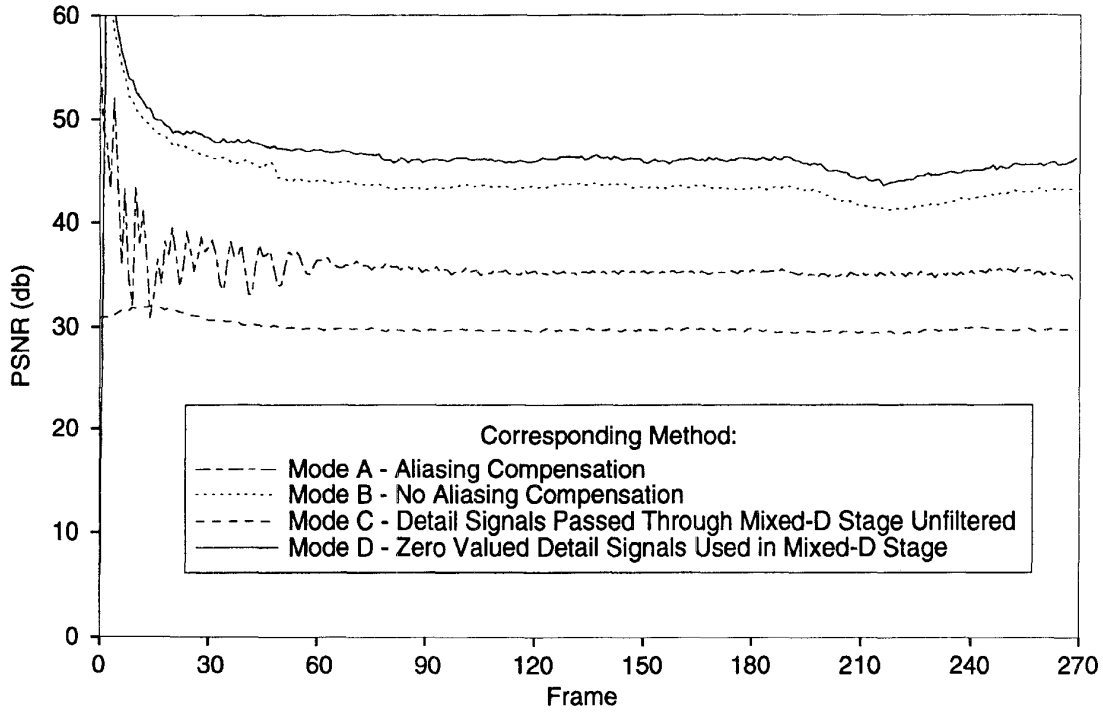

Fig. 7. PSNR values for combined DWT/mixed-D filtering. (PSNR computed relative to mixed-D filtered sequence).

contains one-quarter the number of these unique values;[3] that is, $C_{2^1} = C_{2^0}/4$. Using the results from Section III-C, the forward 2-D DWT therefore requires $3C_{2^1}$ complex additions and C_{2^0} complex multiplications per subband signal, and the inverse 2-D DWT requires C_{2^0} complex multiplications per subband signal. In addition, the 2D DWT^{-1} requires $3C_{2^0}$ complex additions to recombine the upsampled subband signals. Thus, there are a total of $12C_{2^1} + 3C_{2^0}$ complex additions and $8C_{2^0}$ complex multiplications required to the compute the full 2-D DWT.

The filter coefficients $H_l^{2^0}(k_1, k_2)$ each require a total of C_{2^0} complex storage locations and, thus, the total number of complex storage locations required for both the analysis and synthesis filters is $8C_{2^0}$. Since a critically sampled system is used, no additional storage is required for the subband signals themselves. However, additional buffer storage is required after upsampling and before the upsampled subband signals are added together. It is assumed that one of these buffers can also be used to store the combined signals; thus, an additional $3C_{2^0}$ complex storage locations are required during upsampling. Therefore, there are a total of $11C_{2^0}$ additional complex storage locations required for computation of all subband signals.

Since the number of LDE operations is changed by the combined approach, the complexity of LDE filtering must also be considered. For the second-order LDE filters used above, each LDE filter associated with a frequency 2-tuple requires five complex multiplications, four complex additions, and five complex storage locations. In practice, the aliasing-compensated filters have been implemented as four separate LDE filters, one per adjacent quadrant.

Using the above values, the total number of additional arithmetic operations and storage locations for all four test modes are shown in Table I. Negative entries in this table indicate LDE filtering operations no longer performed on the entire signal; next to the negative entries, the corresponding operations required for LDE filtering in the subband signals are provided. It is interesting to note that of the four methods, Mode D, which yielded the highest PSNR values, required the least number of computations and storage and, in fact, required fewer operations and storage than the direct mixed-D filtering method.

V. Conclusion

In this paper, the mixed-D filtering method and the discrete wavelet transform (DWT) have been reviewed. A novel filtering method, which combines these two methods by pre- and postprocessing mixed-D filtering with the DWT, is proposed. It is shown that no additional LDE filters are required in combined DWT/mixed-D filtering, although the frequency-shuffling operation inherent in the DWT requires that the LDE filters be used at different frequency locations than in the mixed-D filtering method. This result indicates that Mixed-D processing is amenable to use in subband coding systems, which are receiving increasing attention in the signal processing literature.

After presenting the details of combined DWT/mixed-D filtering, an example is provided in which a 3-D LT signal is filtered using several modes of the proposed method. The

[3] This is a reasonable assumption, particularly for large-sized signals. The exact distribution of unique values within the subbands depends upon the order of LDE filtering with the subbands.

results from this example indicate that the proposed method can be computationally attractive while providing output sequences that are comparable in quality to those obtained using mixed-D filtering.

Future work in this area might involve the use of finer frequency partitioning, for example, by using nondyadic wavelets or wavelet packets [3]. The use of an adaptive algorithm to determine when subband signals should or should not be mixed-D filtered would also be a possibility for future investigation. Finally, since mixed-D filtering has proven useful in velocity estimation and tracking [9], we can envision a multiresolution velocity estimation system based upon the proposed method.

ACKNOWLEDGMENT

The authors thank N. Bartley for valuable discussions about mixed-D filtering of LT signals. The helpful comments and suggestions of the anonymous reviewers are also acknowledged.

REFERENCES

[1] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 7, pp. 674–693, July 1989.
[2] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909–996, Nov. 1988.
[3] ——, *Ten Lectures on Wavelets*. Philadelphia: SIAM, CMBS Series, 1992.
[4] P. P. Vaidyanathan, *Multirate Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1992.
[5] A. Akansu and R. Haddad, *Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets*. New York: Academic, 1992.
[6] M. S. Lazar and L. T. Bruton, "The design of compactly supported orthonormal wavelets with integer scaling factors," in *Proc. IEEE Int. Symp. Time-Freq. Time-Scale Anal.*, Victoria, Canada, Oct. 1992, pp. 323–326.
[7] L. T. Bruton and N. R. Bartley, "Three-dimensional image processing using the concept of network resonance," *IEEE Trans. Circuits Syst.*, vol. CAS-32, no. 7, pp. 664–672, July 1985.
[8] ——, "Applications of complex filters to realize three-dimensional combined DFT/LDE transfer functions," *IEEE Trans. Circuits Syst.-II: Analog Digital Signal Processing*, vol. 39, no. 6, pp. 391–394, June 1992.
[9] ——, "The enhancement and tracking of moving objects in digital images using adaptive three-dimensional recursive filters," *IEEE Trans. Circuits Syst*, vol. CAS–33, no. 6, pp. 604–612, June 1986.
[10] A. A. Choudhury and L. T. Bruton, "Multidimensional filtering using combined discrete fourier transform and linear difference equation methods," *IEEE Trans. Circuits Syst.*, vol. 37, no. 2, pp. 223–231, Feb. 1990.
[11] K. S. Knudsen and L. T. Bruton, "Mixed domain filtering of multidimensional signals," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 1, no. 3, pp. 260–268, Sept. 1991.
[12] Y. Zhang and S. Zafar, "Motion-compensated wavelet transform coding for color video compression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 2, no. 2, pp. 285–296, Sept. 1992.
[13] M. Vetterli and K. M. Uz, "Multiresolution coding techniques for digital television: A review," *Multidimen. Syst. Signal Processing*, vol. 3, pp. 161–187, 1992.
[14] K. M. Uz, M. Vetterli, and D. J. LeGall, "Interpolative multiresolution coding of advanced television with compatible subchannels," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 1, no. 1, pp. 86–99, Mar. 1991.
[15] G. Karlsson and M. Vetterli, "Extension of finite length signals for sub-band coding," *Signal Processing*, vol. 17, pp. 161–168, 1989.
[16] J. Woods, Ed., *Subband Image Coding*. Boston: Kluwer, 1991.
[17] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
[18] C. S. Burrus and T. W. Parks, *DFT/FFT Convolution Algorithms*. New York: Wiley, 1985.
[19] S. Tong, "Summer project report on 3D DFT/LDE filtering," Univ. Calgary Dep. Elect. Eng, Tech. Rep., July 1990.
[20] A. Kojima, N. Sakurai, and J. Kishigami, "Motion detection using 3D-FFT spectrum," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Minneapolis, MN, Apr. 1993, vol. 5, pp. 213–216.
[21] B. Jahne, *Digital Image Processing*. Berlin: Springer-Verlag, 1991.
[22] M. Lazar, "Applications of multiresolution analysis in multidimensional signal processing," Ph.D. dissertation, Univ. Calgary, 1994.

Michael S. Lazar (M'84) received the B.Sc. degree in computer engineering from the University of Alberta, Canada, in 1984, and the Ph.D. degree in electrical engineering from the University of Calgary, Canada, in 1994.

During 1985–1986 and 1988–1990, he was with Bell Northern Research, and is currently a founding partner of World Web Technologies in Calgary, Alberta, Canada. His research interests include image and video processing and multiresolution analysis.

Leonard T. Bruton (M'71–SM'80–F'81) is a Professor of electrical and computer engineering at the University of Calgary, Alberta, Canada. His research interests are in the areas of analog and digital signal processing. He is particularly interested in the design and implementation of microelectronic digital filters and the applications of multidimensional circuirt and systems theory to digital image processing.