Time-frequency localization in transforms, subbands, and wavelets: a critical review

Richard A. Haddad
Polytechnic University
School of Electrical Engineering and Computer Science
Hawthorne, New York 10532

Ali N. Akansu
Adil Benyassine
New Jersey Institute of Technology
Department of Electrical and Computer Engineering
Center for Communications and Signal Processing Research
University Heights
Newark, New Jersey 07102

Abstract. In many still-image and video processing applications, the time-frequency localization properties of the decomposition technique are an important consideration. While bandwidth compression of the image requires operators with good localization in frequency, spatial features such as edge preservation demand a high degree of localization in time (or the spatial variable). These requirements compete with each other and one is secured at the expense of the other. The classical "uncertainty" principle in the continuous-time domain provides the back drop for this trade-off. Our purpose is to review recent extensions of this principle to the discrete-time case and to develop optimum wave forms. We review common features of block transforms, subband filter banks, and wavelets, and demonstrate how the discrete uncertainty can be used to evaluate these decomposition methods. In particular, we evaluate the trade-off between localization in time and in frequency for several proposed signal decomposition structures.

Subject terms: visual communication; time-frequency localization; block transforms; lapped orthogonal transforms; subband decomposition; wavelet transforms.

Optical Engineering 32(7), 1411–1429 (July 1993).

1 Time-Frequency Distributions and Optimum Signal Shaping

1.1 *Classical Uncertainty*

A basic objective in signal analysis is to devise an operator capable of extracting local features of a signal in both the time and frequency domains. This requires a basis function, or kernel, whose spread or extent is simultaneously narrow in both domains. This in turn suggests that the transformation kernel $\phi(t)$ and its Fourier transform $\Phi(\Omega)$ should have narrow spreads around selected points t_0, and Ω_0.

But the classical uncertainty principle asserts that for any function $\phi(t)$, (with $\sqrt{t}\,\phi(t)\to 0$, as $t\to\pm\infty$),[1-3]

$$\sigma_T \sigma_\Omega \geq \frac{1}{2} , \tag{1}$$

where σ_T and σ_Ω are the rms spread of $\phi(t)$ and $\Phi(\Omega)$ around the center values. That is,

$$\sigma_T^2 = \frac{\int (t-\bar{t})^2 |\phi(t)|^2 \, dt}{E} ,$$

$$\sigma_\Omega^2 = \frac{\frac{1}{2\pi}\int_{-\infty}^{\infty} (\Omega-\bar{\Omega})^2 |\Phi(\Omega)|^2 \, d\Omega}{E} , \tag{2}$$

where E is the energy in the signal.

$$E = \int_{-\infty}^{\infty} |\phi(t)|^2 \, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\Phi(\Omega)|^2 \, d\Omega , \tag{3}$$

and $\bar{t}$ and $\bar{\Omega}$ refer to the center of mass of these kernels,

$$\bar{t} = \frac{\int_{-\infty}^{\infty} t|\phi(t)|^2 \, dt}{E} ,$$

$$\bar{\Omega} = \frac{\frac{1}{2\pi}\int_{-\infty}^{\infty} \Omega|\Phi(\Omega)|^2 \, d\Omega}{E} . \tag{4}$$

The product $\sigma_T \sigma_\Omega$ is called the resolution cell. The equal sign holds in Eq. (1) if and only if $\phi(t)$ (and consequently, its Fourier transform) is Gaussian of the form $\exp(-\alpha t^2)$. The derivation of this result can be found in Papoulis.[2]

The short-time Fourier transform (STFT) has been used to extract frequency characteristics of a signal over some selected interval in time. The STFT positions a window $g(t)$ at some point τ on the time axis and calculates the Fourier transform of the signal contained within the spread or extent of that window,

$$F(\beta,\tau) = \int_{-\infty}^{\infty} f(t) g(t-\tau) \exp(j\beta t) \, dt . \tag{5}$$

When the window $g(t)$ is Gaussian, the STFT is called a Gabor transform.[3-5] The basis functions are generated by *modulation* and *translation* of the window function by pa-

Paper VCI-05 received Dec. 8, 1992; revised manuscript received Feb. 22, 1993; accepted for publication Feb. 24, 1993.

rameters β and τ, respectively. Note that when τ increases, the kernel simply translates in time while keeping the spread of the window fixed.

Let $g(t) \leftrightarrow G(\Omega)$ be a Fourier transform pair, and assume that $\bar{t}=0$ and $\bar{\Omega}=0$. Then the translated, modulated kernel pair are given by

$$g_{\tau,\beta}(t) = g(t-\tau)\ \exp(j\beta t) \leftrightarrow G_{\tau,\beta}(\Omega)$$
$$= \exp[-j(\Omega-\beta)\tau]G(\Omega-\beta)\ . \tag{6}$$

This two-parameter family is centered at τ,β in the time-frequency plane, i.e.,

$$\begin{aligned} \bar{t}_{\tau,\beta} &= \tau, \\ \bar{\Omega}_{\tau,\beta} &= \beta\ . \end{aligned} \tag{7}$$

Now it is readily shown that the spread of this shifted, modulated kernel is constant in both domains, i.e.,

$$\sigma^2_{\Omega(\tau,\beta)} = \frac{\frac{1}{2\pi}\int_{-\infty}^{\infty}(\Omega-\beta)|G_{\tau,\beta}(\Omega)|^2\ d\Omega}{E} = \sigma_\Omega^2\ ,$$

$$\sigma^2_{\tau,\beta)} = \int_{-\infty}^{\infty}(t-\tau)^2|g_{\tau,\beta}(t)|^2\ dt = \sigma_T^2\ , \tag{8}$$

where σ_Ω and σ_T are the rms spreads of the unmodulated, untranslated kernels. This implies that the resolution cell $\sigma_T\sigma_\Omega$ has a constant shape as well as a constant area in the time-frequency plane, as shown in Fig. 1. For the Gaussian window, $\sigma_T\sigma_\Omega = 1/2$. The wavelet transform introduced in Sec. 5 has a variable shape but a constant product.

1.2 *Discrete-Time Uncertainty*

The discrete-time version is as follows: Let $f(n) \leftrightarrow F[\exp(j\omega)]$ be a discrete-time Fourier transform pair,

$$F[\exp(j\omega)] = \sum_{n=-\infty}^{\infty} f(n)\ \exp(-jn\omega) \leftrightarrow f(n)$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} F[\exp(j\omega)]\ \exp(jn\omega)\ d\omega\ . \tag{9}$$

By the Parseval's theorem, the energy is

$$E = \sum_{n=-\infty}^{\infty}|f(n)|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|F[\exp(j\omega)]|^2\ d\omega\ . \tag{10}$$

We define the mean (analogous to the center of mass of a distribution) by[6-8]:

$$\bar{\omega} = \frac{\frac{1}{2\pi}\int_{-\pi}^{\pi}\omega|F[\exp(j\omega)]|^2\ d\omega}{E}\ , \tag{11}$$

$$\bar{n} = \frac{\sum_{-\infty}^{\infty} n|f(n)|^2}{E}\ . \tag{12}$$

The spread of a function in time and in frequency is

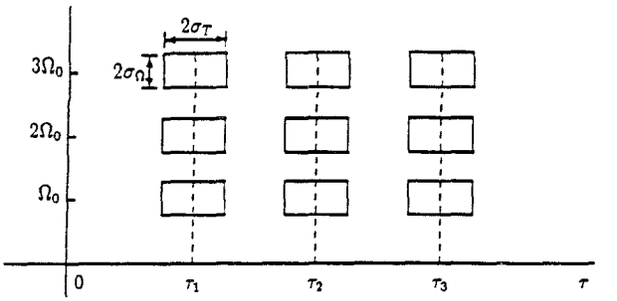


Fig. 1 Time-frequency plane showing resolution cells for STFT.

$$\sigma_n^2 = \frac{\sum_{-\infty}^{\infty}(n-\bar{n})^2|f(n)|^2}{E}\ ,$$

$$\sigma_\omega^2 = \frac{\frac{1}{2\pi}\int_{-\pi}^{\pi}(\omega-\bar{\omega})^2|F[\exp(j\omega)]|^2\ d\omega}{E}\ . \tag{13}$$

For any real signal $\bar{\omega}=0$, and without loss of generality, we can also shift the time origin to make $\bar{n}=0$. For this case, in Sec. 7.1 we show that the time-frequency product $\sigma_n\sigma_\omega$ or resolution cell is given by

$$\begin{aligned} \sigma_n\sigma_\omega &\geq \frac{|1-\mu|}{2} \\ \mu &\triangleq \frac{|F[\exp(j\omega)]_{\omega=\pi}|^2}{E} = \frac{|F(-1)|^2}{E}\ . \end{aligned} \tag{14}$$

In the analog version, $F(\pm\infty)=0$ and the lower limit is simply $1/2$. In the discrete-time case $F(-1)$ need not be zero. Note that in our notation, $F[\exp(j\omega)]$ at $\omega=0$ and $\omega=\pi$ are denoted by $F(1)$ and $F(-1)$, respectively.

Remark. The frequency measure in Eq. (13) is not suitable for bandpass signals with peak frequency responses centered at $\pm\hat{\omega}$. To obtain a measure of the spread about $\hat{\omega}$, we need to define σ_ω^2 on the interval $[0,\pi]$, rather than $[-\pi,\pi]$. In this case we use

$$\hat{\omega} = \frac{\frac{1}{\pi}\int_0^{\pi}\omega|F[\exp(j\omega)]|^2\ d\omega}{\frac{1}{\pi}\int_0^{\pi}|F[\exp(j\omega)]|^2\ d\omega}\ , \tag{15}$$

$$\hat{\sigma}_\omega^2 = \frac{\frac{1}{\pi}\int_0^{\pi}(\omega-\hat{\omega})^2|F[\exp(j\omega)]|^2\ d\omega}{\frac{1}{\pi}\int_0^{\pi}|F[\exp(j\omega)]|^2\ d\omega}\ , \tag{16}$$

and σ_n^2 remains unchanged. It easily follows that

$$\hat{\sigma}_\omega^2 = \sigma_\omega^2 - (\hat{\omega})^2$$

and

$$\hat{\sigma}_\omega^2\sigma_n^2 \geq \frac{1}{4}(1-\mu)^2 - (\hat{\omega})^2\sigma_n^2\ . \tag{17}$$

Equation (17) demonstrates the reduction in the time-frequency product when using the $[0,\pi]$ interval for bandpass signals. An alternative derivation similar to Sec. 7 shows that this product can be expressed as

$$\hat{\sigma}_\omega \sigma_n \geq \frac{1}{2}|1-\mu'| \ ,$$

$$\mu' = \frac{\hat{\omega}}{\pi}\frac{|F(1)|^2}{E} + \left(1-\frac{\hat{\omega}}{\pi}\right)\frac{|F(-1)|^2}{E} \ . \tag{18}$$

For bandpass signals with zero dc gain, $F(1)=0$, Eq. (18) reduces to

$$\mu' = \left(1-\frac{\hat{\omega}}{\pi}\right)\frac{|F(-1)|^2}{E} \ .$$

Additionally, if we have $F(-1)=0$, then $\mu'=0$ and

$$\hat{\sigma}_\omega \sigma_n \geq \frac{1}{2} \ .$$

In the sequel, we concentrate on low-pass filters such that $F[\exp(j\omega)]_{\max}$ occurs at $\omega=0$, and Eqs. (13) and (14) are used. In this case, there are two classes of filters or signals:

$$\text{class I:} \quad F(-1)=0 \rightarrow \sigma_n\sigma_\omega \geq \frac{1}{2} \ , \tag{19}$$

$$\text{class II:} \quad F(-1)\neq 0 \rightarrow \sigma_n\sigma_\omega \geq \frac{|1-\mu|}{2} \ . \tag{20}$$

The bound on the time-frequency product in the first case is the same as that for the continuous-time case (in which $F(\pm\infty)=0$). In the analog case, we know that the equality in the lower bound is achieved when $\Omega F(\Omega)$ is proportional to $dF/d\Omega$, or $F(\Omega)=K\exp(-b\Omega^2/2)$, a Gaussian. In the discrete-time formulation in Sec. 7, we have the same form of integral resulting in the differential equation,

$$\frac{dF}{d\omega} = -K\omega F^*[\exp(j\omega)] \ ,$$

whose solution is a Gaussian $\exp(-K\omega^2/2)$. This Gaussian function satisfies the differential equation but cannot satisfy the class I boundary condition $F[\exp(j\pi)]=0$. In this case, we conclude that the lower bound cannot be attained and the strict inequality holds, $\sigma_n\sigma_\omega > 1/2$.

We show that the binomial function is a finite impulse response (FIR) approximation to the Gaussian that matches the zero boundary condition at $\omega=\pi$.

For the class II set of functions, the Gaussian can satisfy both the differential equation and the boundary condition resulting in the equality $\sigma_n\sigma_\omega = 1/2|1-\mu|$.

1.3 *Gaussian and Binomial Distributions*

For the class II signals, the Gaussian is

$$|F[\exp(j\omega)]|^2 = K\exp(-\omega^2/2\sigma^2), \qquad |\omega|<\pi \tag{21}$$

$$K = \frac{(\pi/2)^{1/2}}{\sigma \operatorname{erf}(\pi/\sigma)} \ .$$

The constant K is chosen to normalize the energy E to unity over $[-\pi,\pi]$. In Sec. 7.2, we show that

$$\sigma_\omega = \sigma(1-\mu)^{1/2} \ , \qquad \mu = \frac{|F(\pi)|^2}{E} = K\exp(\pi^2/2\sigma^2) \ , \tag{22}$$

and, hence,

$$\sigma_\omega\sigma_n = \frac{|1-\mu|}{2} \ , \rightarrow \sigma_n = \frac{(1-\mu)^{1/2}}{2\sigma} \ . \tag{23}$$

For the narrow-band case, $\sigma \leq \pi/4$, $\mu \leq 10^{-3}$, and $F(-1) \approx 0$, resulting in $\sigma_\omega^2 \approx \sigma^2$, $\sigma_\omega\sigma_n \approx 1/2$. The corresponding time function is found to be approximately Gaussian:

$$f(n) = \sigma(K/\pi)^{1/2}\exp(-\sigma^2 n^2) \ . \tag{24}$$

Examples of these narrow-band Gaussian functions are shown in Fig. 2. Again note that the time-frequency product is very close to 1/2 in these cases.

For the wide-band case, with $\sigma \geq 3\pi/8$, we must use the more exact expansions in Eqs. (22) and (23). For example, for $\sigma=\pi/2$, we calculate $\mu=0.22625$, $\sigma_\omega=1.382$, $\sigma_n\sigma_\omega=0.3869$, and $\sigma_n=0.28$ time samples. In this case, there is no simple approximation for $f(n)$ that must be computed numerically from the inversion formula, Eq. (9). These are shown in Fig. 3 in which the very short duration of $f(n)$ is duly noted.

The binomial sequences are a family generated by successive differences of the binomial kernel,[9] as summarized in Table 1.

The kernel is the binomial sequence $\binom{N}{k}$ that looks similar to a sampled truncated Gaussian in time, and the frequency response resembles the Gaussian $\exp(-\omega^2/2\sigma^2)$. To demonstrate this, let us compare $f(\omega)=[\cos(\omega/2)]^N$, with $g(\omega)=\exp(-\omega^2/4\sigma^2)$. Taking logarithms,

$$\ln[g(\omega)] = -\frac{\omega^2}{4\sigma^2} \ ,$$

$$\ln[f(\omega)] = -N\log\cos\frac{\omega}{2} = N\left[-\frac{1}{2}\left(\frac{\omega}{2}\right)^2 - \frac{1}{12}\left(\frac{\omega}{2}\right)^4 \ldots\right] \ .$$

Matching quadratic terms, $N=2/\sigma^2$ in these expansions results in a normalized error that is

$$\frac{g(\omega)-f(\omega)}{g(\omega)} \approx \frac{(\omega/\sigma)^4}{48N} \ , \qquad \frac{\omega}{\sigma} < 2 \ .$$

The localization features of the binomial kernel are as follows: $X(1)=0$, so that $\sigma_n\sigma_\omega > 1/2$. From Sec. 7.3,

$$E = \sum_{k=0}^{N}\binom{N}{k}^2 = 2^{2N}\frac{[1.3.5\ldots(2N-1)]}{(2.4.6\ldots 2N)} \ , \tag{25}$$

$$\bar{n} = N/2 \ , \tag{26}$$

$$\sigma_n = \frac{N}{2(2N-1)^{1/2}} \ . \tag{27}$$

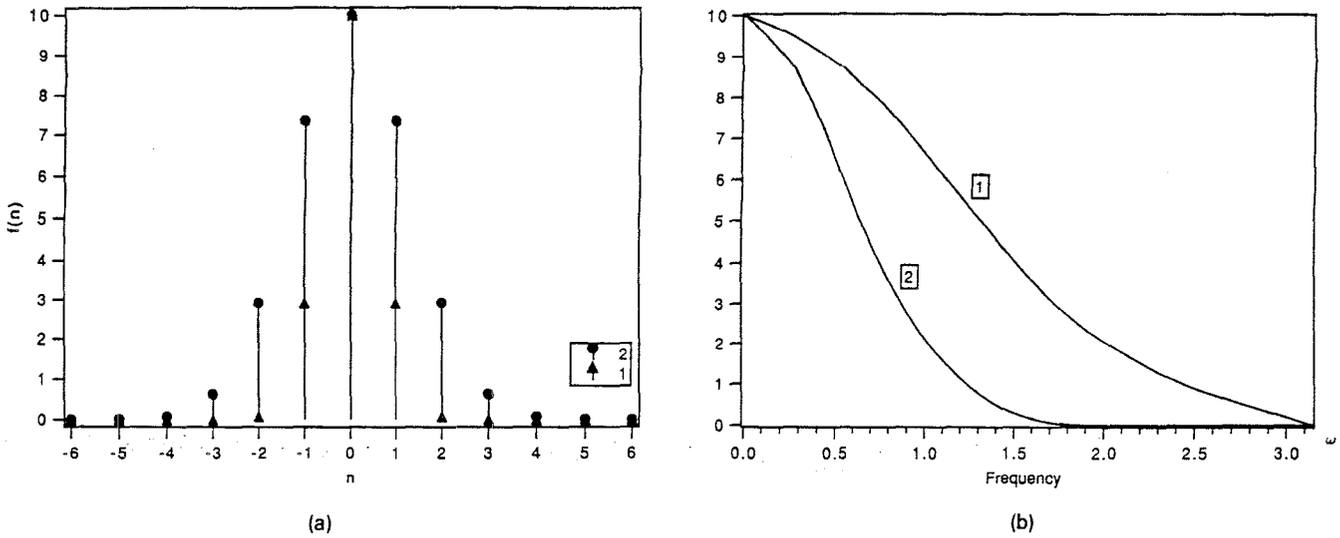


Fig. 2 (a) Time and (b) frequency plots for narrow-band Gaussian functions: (1) $\sigma_\omega = \pi/4$, $\sigma_n = 0.637$ samples and (2) $\sigma_\omega = \pi/8$, $\sigma_n = 1.274$ samples.

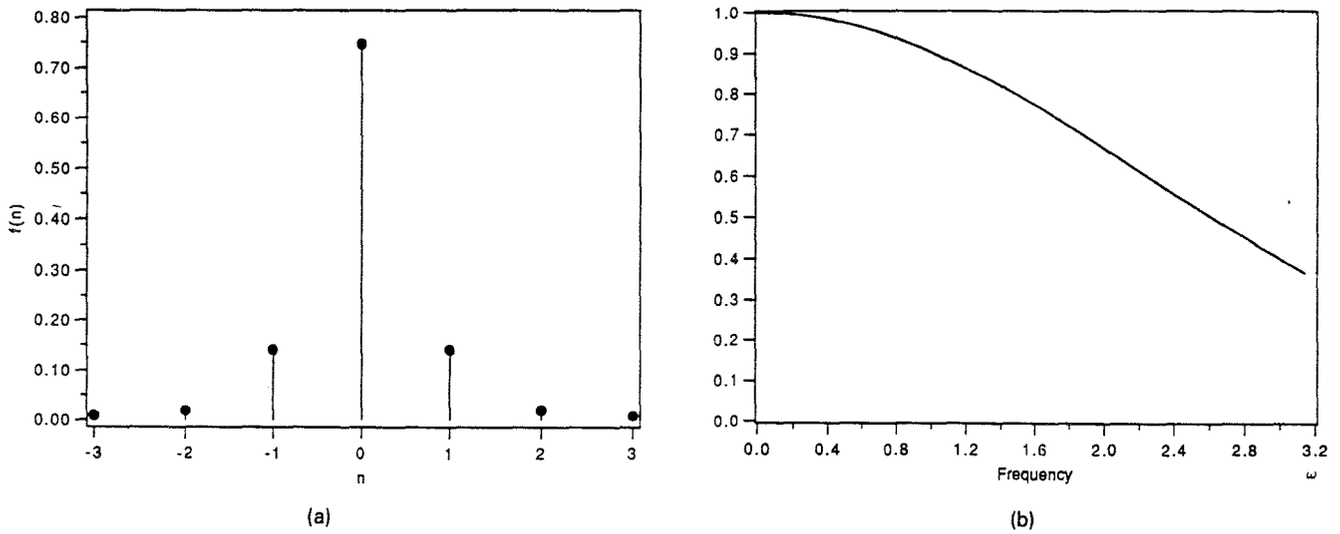


Fig. 3 (a) Time and (b) frequency plots for the wideband Gaussian case: $\sigma_\omega = \pi/2$ and $\sigma_\omega \sigma_n = 0.3869$.

Table 1 The binomial family: $H_r(k)$ are discrete Hermite polynomials.

Time Function	⟷	Transform
$x_0(k) = \binom{N}{k}, \quad 0 \le k \le N$		$X_0(z) = (1+z^{-1})^N$ $\lvert X_0(e^{j\omega})\rvert = 2^N(\cos\frac{\omega}{2})^N$
$x_r(k) = \nabla^r \binom{N-r}{k}$ $x_r(k) = H_r(k)\binom{N}{k}$		$X_r(z) = (1+z^{-1})^{N-r}(1-z^{-1})^r$ $\lvert X_r(e^{j\omega})\rvert = 2^N(\cos\frac{\omega}{2})^{N-r}(\sin\frac{\omega}{2})^r$

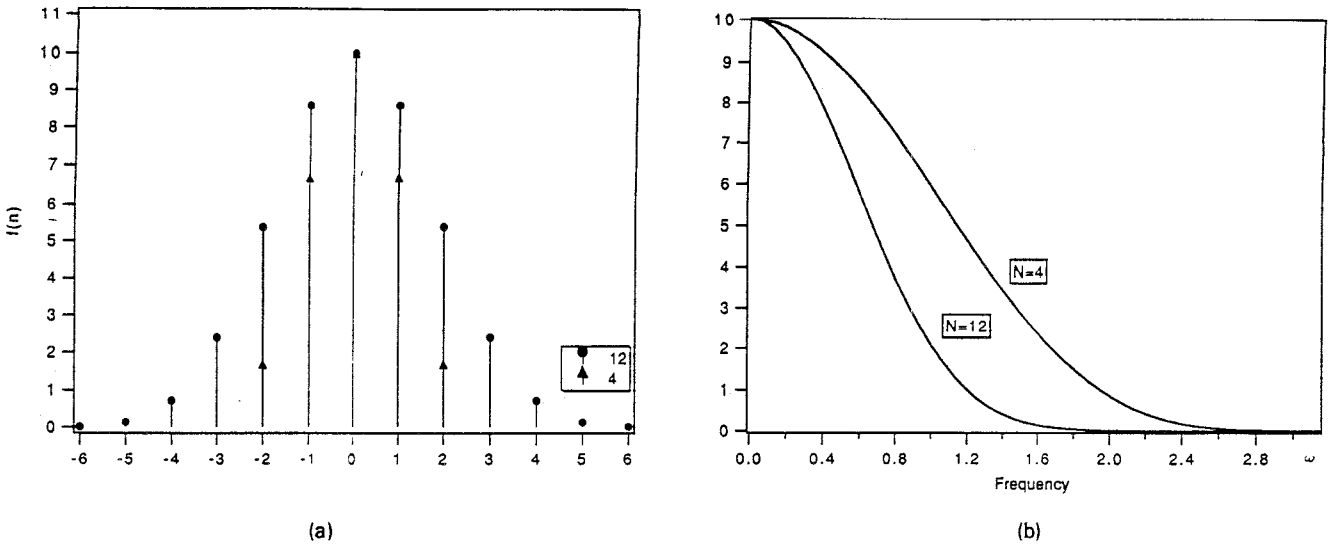


Fig. 4 Time-frequency plots of binomial sequences: (a) $N=4(\sigma_\omega=0.665,\ \sigma_n=0.756,$ and $\sigma_\omega\sigma_n=0.50274)$ and (b) $N=12(\sigma_\omega=0.4,\ \sigma_n=1.25,$ and $\sigma_n\sigma_\omega=0.5002)$.

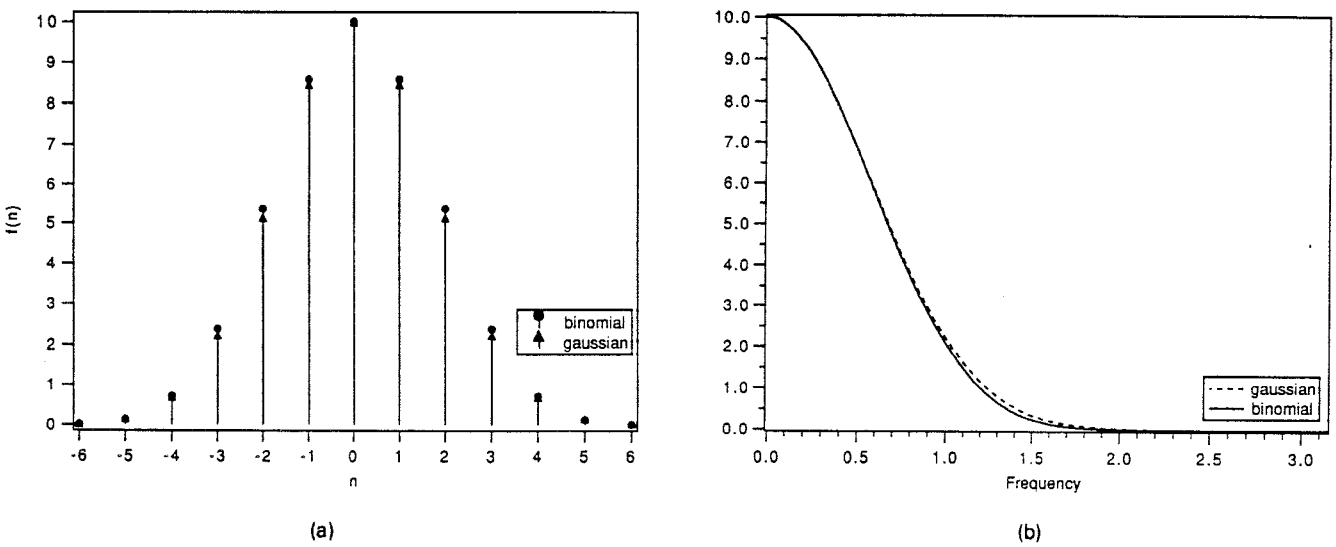


Fig. 5 Binomial and Gaussian (a) time and (b) frequency plots for $N=12$ and $\sigma=1/\sqrt{6}$, respectively.

It can also be shown that

$$\sigma_\omega^2=\frac{4(A+B)}{E}$$

$$A=\sum_{k=0}^{N-1}\binom{2N}{k}\frac{(-1)^{N-k}}{(N-k)^2}, \tag{28}$$

$$B=\binom{2N}{N}\frac{\pi^2}{12}.$$

Sample binomial time-frequency responses are displayed in Fig. 4. Note that these approximate the Gaussian very well, and the time-frequency products are, respectively, 0.50274 and 0.5002, for $N=4.12$.

Figure 5 shows binomial and Gaussian responses on the same axes for $N=12$ and $\sigma=\sqrt{2/N}=1/\sqrt{6}$. Both time and frequency plots are almost indistinguishable. We conclude that the binomial filter provides a simple yet excellent FIR approximation to the optimum Gaussian wave form in the time and frequency domains.

1.4 *Band-Pass Filters*

The binomial family can also provide good approximations to Gaussian bandpass filters. The r'th member $X_r(z)$ has a magnitude square response of the form

$$\left(\sin\frac{\omega}{2}\right)^{2r}\left(\cos\frac{\omega}{2}\right)^{2(N-r)}.$$

For N large, this response is approximately even symmetric around $\hat{\omega}\approx 2\sin^{-1}(r/N)^{1/2}$, and approximately Gaussian. A major advantage of these binomial filters is that they can be synthesized using only add, subtract, and delay operators.[9]

A second class of bandpass filters can be obtained by modulating the low-pass binomial. Let $h(n) \triangleq x_0(n)$, and let

$$g(n)=\frac{1}{\sqrt{2}}h(n)\ \cos\omega_0 n \leftrightarrow G[\exp(j\omega)]$$

$$=\frac{1}{\sqrt{2}}(H\{\exp[j(\omega-\omega_0)]\}+H\{\exp[j(\omega+\omega_0)]\}) \quad . \quad (29)$$

Now if $\hat{\omega}=\omega_0 >> \sigma = 2/\sqrt{N}$, the leakage of the tail of $H\{\exp[j(\omega+\omega_0)]\}$ into the frequency band near $\omega=\omega_0$ is negligible, and over $[0,\pi]$,

$$G[\exp(j\omega)]=\frac{1}{\sqrt{2}}H\{\exp[j(\omega-\omega_0)]\}$$

$$=\frac{1}{\sqrt{2}}X_0\{\exp[j(\omega-\omega_0)]\} \quad . \quad (30)$$

In this case, it is easy to show that the spread in time and frequency domains is the same as for the low-pass binomial prototype window.

2. Time-Frequency Properties of Block Transforms

2.1 *Orthonormal Sequences and Block Transforms*

Consider a set of N sequences $\{\phi_r(k)\}$, $r=0, 1, ..., N-1$ on the interval $[0, N-1]$. These are orthonormal if

$$\sum_{n=0}^{N-1}\phi_r(n)\phi_s^*(n)=\delta_{r-s} \quad . \quad (31)$$

Then, any sequence $f(n)$ on $[0, N-1]$ can be represented by the orthonormal expansion

$$f(k)=\sum_{r=0}^{N-1}\theta_r\phi_r(k) \quad , \qquad 0\leq k\leq N-1 \quad ,$$

where

$$\theta_r=\sum_{k=0}^{N-1}f(k)\phi_r^*(k) \quad , \qquad 0\leq r\leq N-1 \quad (32)$$

θ: the spectral coefficients.

From the foregoing, one can show that the Parseval's theorem holds,

$$E=\sum_{k=0}^{N-1}|f(k)|^2=\sum_{r=0}^{N-1}|\theta_r|^2 \quad . \quad (33)$$

The orthogonality is also demonstrated in the frequency domain via

$$\sum_k\phi_r(k)\phi_s^*(k)=$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_r[\exp(j\omega)]\Phi_s^*[\exp(j\omega)]\ d\omega=\delta_{r-s} \quad . \quad (34)$$

To express the foregoing in the format of block transforms, let the data set, $\{f_0, f_1, ..., f_{N-1}\}$, and the spectral coefficients, $\{\theta_0, \theta_1, ..., \theta_{N-1}\}$, be designated by row vectors $\mathbf{f}^T$ and $\boldsymbol{\theta}^T$, respectively. Furthermore, let $[\phi_r(k),\ 0\leq k\leq N-1]$ be the r'th row of the matrix $\Phi=[\phi_r(k)]$. Then Eqs. (32) and (33) take the form

$$\boldsymbol{\theta}=\boldsymbol{\Phi}^*\mathbf{f} \quad ,$$

$$\mathbf{f}=\boldsymbol{\Phi}^T\boldsymbol{\theta} \quad , \quad (35)$$

and

$$\boldsymbol{\theta}^T\boldsymbol{\theta}^*=\mathbf{f}^T\mathbf{f}^* \quad . \quad (36)$$

A matrix with the property $\Phi^*\Phi^T=I$ is said to be unitary.

The spectral coefficient vector $\boldsymbol{\theta}$ can be encoded, quantized, and transmitted as suggested by Fig. 6(a). In the absence of quantization and transmission errors, $\hat{\mathbf{f}}=\mathbf{f}$. These spectral coefficients can also be generated by the serial feed filter bank structure of Fig. 6(b). The analysis bank of FIR filters has impulse responses $h_r(n)$ that are time-reversed and translated basis sequences $\phi_r^*(N-1-n)$. The downsampling of every N cycles generates the coefficient vector $\boldsymbol{\theta}$ from the previous batch of N data samples.

In the synthesis section, the coefficients $\{\theta_i\}$ are interlaced with zeros by the upsampler and filtered by a bank of FIR synthesis/interpolation filters with impulse responses $g_r(n)=\phi_r(n)$. In the absence of quantization and transmission errors, $\hat{f}(n)$ is simply a delayed version, $f[n-(N-1)]$, of the input sequence.

Yet another view is possible if we convert the serial feed of Fig. 6(b) into a parallel feed (as by a buffer) as shown in Fig. 6(c), and replace the filter bank by the unitary matrix Φ^*. Hence, we have three equivalent representations of block transforms, each providing insight into the transformations and the multirate filter banks.

2.2 *Time-Frequency Behavior*

We can examine the time-frequency localization of established block transforms by analyzing the time-frequency products of each synthesis filter. Figure 7 shows the time-frequency wave forms of the 8×8 discrete cosine transform (DCT) and Walsh-Hadamard Transform (WHT) basis functions. The time-frequency spreads for the DCT and WHT are given in Table 2. The trade-offs in σ_n and σ_ω as a function of filter length are obvious from this table. To sharpen the frequency response, the transform size or filter lengths are increased. As seen in Table 2, the frequency spread σ_ω^2 is decreased significantly but at an appreciable increase in σ_n^2, owing to the longer filter lengths or basis sequences. The time-frequency products or resolution cells also increase with filter length, i.e., with size of transform.

The representations in Fig. 6 show that for block transforms, the filter lengths are equal to the number of filters that correspond to the size of the transforms. There is a very limited flexibility. For compression purposes, narrow frequency bands are desirable, and we are led to consider a broader structure, the M-band filter bank, where the length of each analysis and synthesis filter is not constrained by the number of filters. These extended length or overlapping basis filters also provide additional degrees of freedom for optimizing other aspects of system performance.

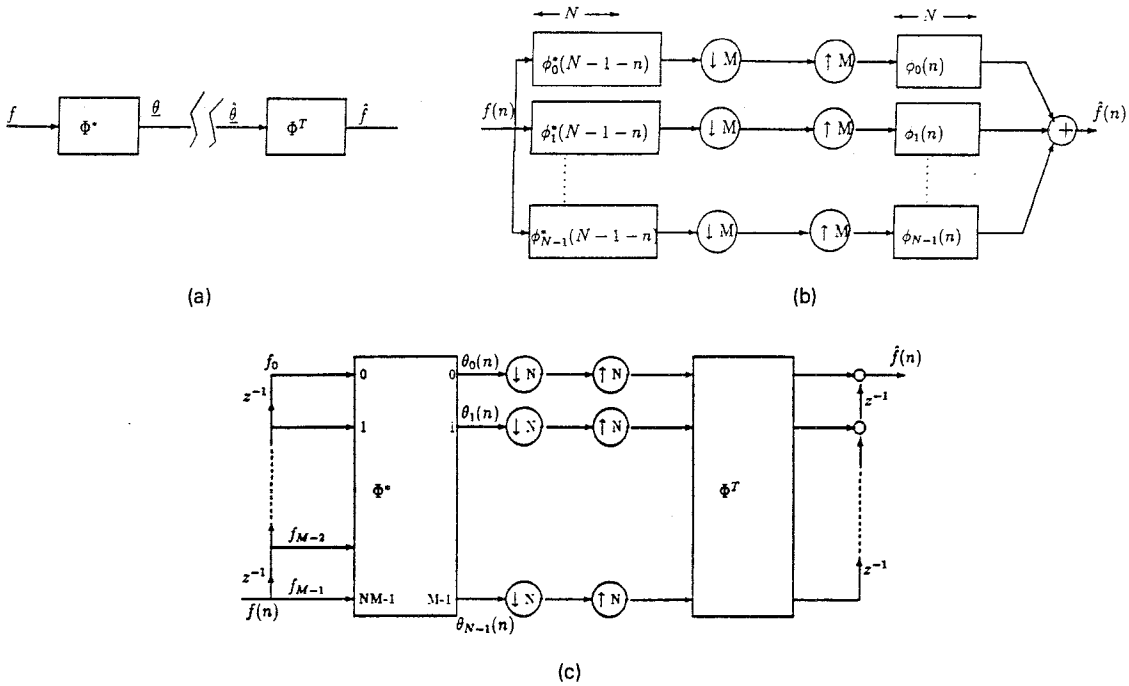


Fig. 6 Unitary transforms represented by (a) block feed structure, (b) serial feed filter bank realization, and (c) alternate interpretation of block transforms.

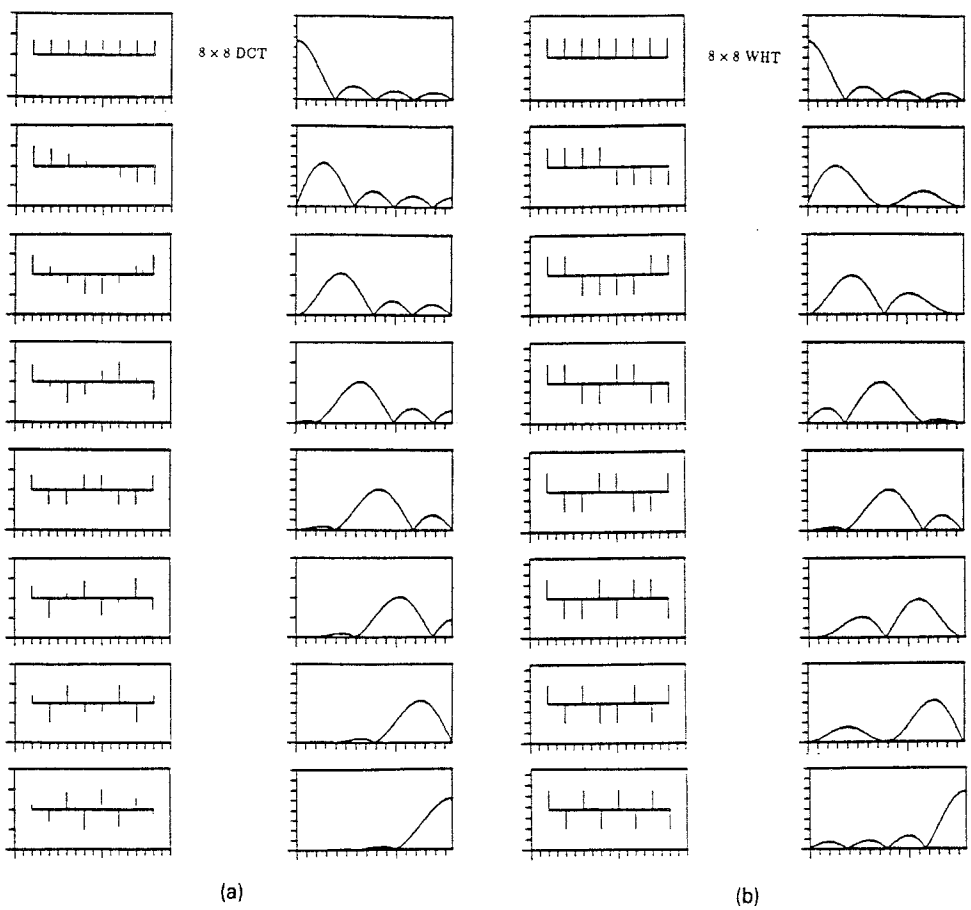


Fig. 7 The basis functions of (a) 8 × 8 DCT and (b) 8 × 8 WHT in time and frequency domains.

Table 2 Time-frequency localizations of DCT and WHT bases for two-, four-, and eight-band cases.[8]

	$\bar{\omega}$	$\bar{n}$	σ_ω^2	σ_n^2	$\sigma_\omega^2 \times \sigma_n^2$
2 × 2 DCT	0	0.50	1.2899	0.2500	0.3225
and WHT	π	0.50	1.2899	0.2500	0.3225
4 × 4 DCT	0	1.50	0.6787	1.2500	1.2234
	1.27	1.50	0.3809	1.9570	0.7454
	1.85	1.50	0.2424	1.2500	0.3030
	π	1.50	0.4896	0.5428	0.2657
4 × 4 WHT	0	1.50	0.6787	1.2500	0.8484
	1.29	1.50	0.2424	1.2500	0.3030
	1.85	1.50	0.2424	1.2500	0.3030
	π	1.50	0.6787	1.2500	0.8484
8 × 8 DCT	0	3.50	0.3447	5.2500	1.8097
	0.74	3.50	0.3021	8.4054	2.5393
	1.02	3.50	0.2413	5.9572	1.4375
	1.36	3.50	0.1957	5.4736	1.0712
	1.71	3.50	0.1488	5.2500	0.7812
	2.08	3.50	0.1206	5.0263	0.6062
	2.45	3.50	0.0797	4.5428	0.3621
	π	3.50	0.1388	2.0955	0.2908
8 × 8 WHT	0	3.50	0.3447	5.2500	1.8097
	0.82	3.50	0.3485	5.2500	1.8296
	1.15	3.50	0.2977	5.2500	1.5629
	1.43	3.50	0.1488	5.2500	0.7812
	1.72	3.50	0.1488	5.2500	0.7812
	1.99	3.50	0.2977	5.2500	1.5629
	2.33	3.50	0.3485	5.2500	1.8296
	π	3.50	0.3447	5.2500	1.8097

3 M-Band Filter Banks

To obtain narrower frequency bands σ_ω^2, we can expand the length of each analysis and synthesis filter as shown in Fig. 8. We should expect a concomitant increase in σ_n^2. The extension of the unitary block transform, nonoverlapping basis, results to this case is called the paraunitary solution.[10–15] The key results are as follows: the system shown in Fig. 8 is a perfect reconstruction, i.e., $\hat{f}(n) = f(n - n_0)$, if

1. the analysis and synthesis filters have the same length N and

$$G_r(z) = z^{-(N-1)} H_r(z^{-1}) \ ;$$

2. each unit-sample response is normalized to unity,

$$\sum_{n=0}^{N-1} h_r^2(n) = 1 \ ;$$

3. each filter is orthogonal to its own translates shifted by M and its integer multiples, i.e.,

$$\sum_k h_r(k) h_r(Mn + k) = 0 \ , \qquad n \neq 0 \ ;$$

4. filter $h_r(k)$ is orthogonal to $h_s(k)$ and to its translates,

$$\sum_k h_r(k) h_s(Mn + k) = 0 \ .$$

All these can be summarized succinctly by defining the cross correlation,

$$\rho_{rs}(n) = h_r(-n) * h_s(n) \leftrightarrow \Phi_{rs}(z) = H_r(z^{-1}) H_s(z) \ , \qquad (37)$$

and requiring that

$$\rho_{rs}(Mn) = \begin{cases} 0, & r \neq s \\ \delta(n), & r = s \end{cases} \ . \qquad (38)$$

Equation (38) reduces to the simpler orthonormal conditions of Eq. (34) when $N = M$.

The derivation of this result is developed in Ref. 13. For the paraunitary solution, each filter must satisfy $\rho_{rr}(Mn) = \delta(n)$. Another way of saying this is $H_r(z) H_r(z^{-1})$ is an M'th band filter.[13]

As an extension of the unitary matrix approach, the serial feed M-band filter bank can be converted into a block filter bank shown in Fig. 8(b). In that diagram, the input $\mathbf{x}$ is a vector consisting of N blocks of M points each. The symbol $\mathbf{P}^T$ is an $M \times NM$ matrix, the r'th row of which is $h_r(N - 1 - n)$. Similarly, $\mathbf{Q}$ is an $NM \times M$ matrix; the r'th column of $\mathbf{Q}$ corresponds to $h_r(n)$.

The orthogonality in the frequency domain is evident from

$$\rho_{rs}(Mn) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_r[\exp(-j\omega)] H_s[\exp(j\omega)] \exp(jMn) d\omega \ ,$$

$$= \begin{cases} 0, & r \neq s \\ \delta(n), & r = s \end{cases} \ . \qquad (39)$$

The lapped orthogonal transform LOT is a special case[13,14,16–18] of the foregoing when $N = 2M$. We can now compute the time-frequency localization of LOT filter banks and compare them with their block transform antecedents. Table 3 displays the localization characteristics of the 8×8 DCT and the 8×16 DCT-based LOT. Again we note the narrowing of the frequency spread and the attendant increase in σ_n^2. The time-frequency products of the DCT-LOT basis is significantly less than that of DCT, as expected.

4 Hierarchical Filter Banks Based on Two-Band Perfect Reconstruction Quadrature Mirror Filter

The special case of $M = 2$ defines the two-band filter bank; the associated paraunitary conditions are simply

$$\rho_{rs}(2n) = \begin{cases} \delta(n), & r = s \\ 0, & r \neq s, \qquad 0 \leq r, s \leq 1 \end{cases} \ . \qquad (40)$$

For this case, we first design $h_0(n)$ to satisfy $\rho_{00}(2n) = \delta(n)$, i.e., $H_0(z)$ is a spectral factor of the half-band filter $H_0(z) H_0(z^{-1})$. For an even number of taps, this gives $N/2$ nonlinear algebraic equations in N unknowns. The remaining conditions (or degrees of freedom) can be obtained by imposing other requirements on $H(z)$,[19–21] the most notable of which is $H(-1) = 0$. Interestingly enough, as we show later, this condition along with Eq. (40) suffices to define generating functions or interscale coefficients for orthonormal wavelets.[13,22,23]

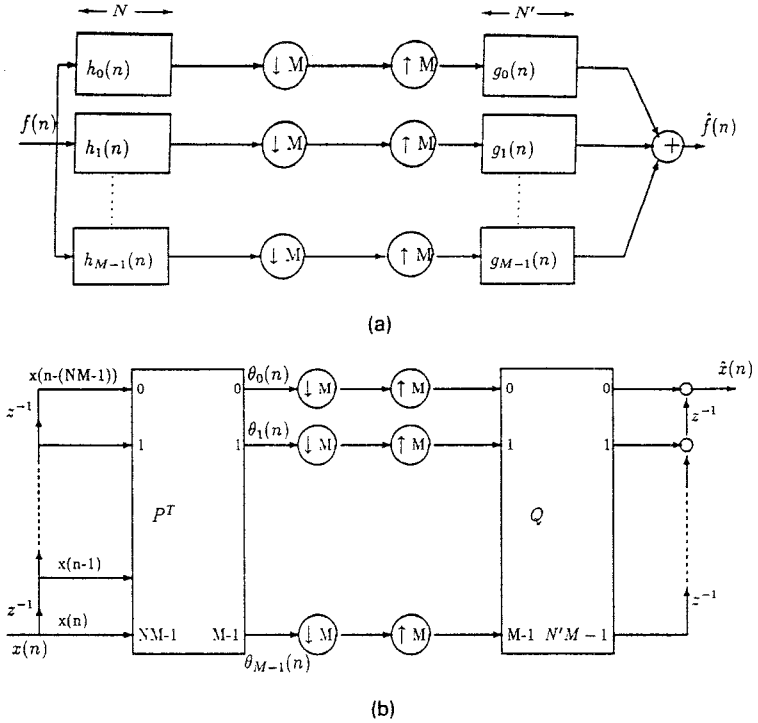


Fig. 8 (a) Critically sampled (maximally decimated) M-band filter bank, $N=N'$ in the paraunitary case and (b) equivalent parallel feed structure.

Having designed $h_0(n)$ to satisfy Eq. (40) and other imposed conditions, we can then obtain the remaining paraunitary analysis and synthesis filters from

$$H_1(z)=z^{-(N-1)}H_0(-z^{-1}) \ ,$$

$$G_0(z)=-H_1(-z) \ , \tag{41}$$

$$G_1(z)=H_0(-z) \ .$$

This two-band decomposition is extended to the binary subband tree structure shown in Fig. 9(a) to obtain a hierarchical M-band structure.[24]

This decomposition into an L level tree results in $M=2^L$ equal-width subbands in the ideal case. The data rate is reduced by a factor of 2 at each level. In this example there are four channels each operating at a data rate of $f_s/4$, resulting in a conservation of the data rate from a full-band input to a multichannel output. This is another example of a critically sampled filter bank. The synthesis structure is simply the inverse of the analysis configuration, and the structure is a perfect reconstruction (PR) if the progenitor two-band filter bank is a PR. By use of the equivalence suggested in Fig. 9(b), the two-level, binary tree is equivalent to the single-level, direct, four-band structure of Fig. 9(c). But note that the latter is far more constrained than a freely chosen four-band filter bank. For example, a nontrivial two-band paraunitary module cannot have linear phase; the unconstrained four-band bank can be both paraunitary and linear phase.

The dyadic or octave-band tree generates naturally a multiresolution signal decomposition. It splits only the lower half of the spectrum into two equal bands at any level of the tree. The detail or high-frequency half-band spectral component of the signal at any level is decomposed no further. Figures 10(a) and 10(b) show the dyadic analysis-synthesis tree for $L=3$ and the equivalent direct structure for the analysis stage, respectively. Figures 10(c) and 10(d) display the schematic of a four-band dyadic tree and the corresponding octave band split, respectively. Note that this structure is also critically sampled. At each level, the signal is decomposed into a coarse approximation at the low band and a detail component at the high-signal band at that resolution or scale. Hence, the designation, multiresolution. This dyadic tree is also the basic structure for the wavelet multiresolution configuration and fast wavelet algorithm.

Table 3 Time-frequency localizations of 8 × 8 DCT and 8 × 16 DCT LOT.[11,18]

	$\bar{\omega}$	$\bar{n}$	σ_ω^2	σ_n^2	$\sigma_\omega^2 \times \sigma_n^2$
8 × 8 DCT	0	3.50	0.3447	5.2500	1.8097
	0.74	3.50	0.3021	8.4054	2.5393
	1.02	3.50	0.2413	5.9572	1.4375
	1.36	3.50	0.1957	5.4736	1.0712
	1.71	3.50	0.1488	5.2500	0.7812
	2.08	3.50	0.1206	5.0263	0.6062
	2.45	3.50	0.0797	4.5428	0.3621
	π	3.50	0.1388	2.0955	0.2908
8 × 16 DCT-LOT	0	7.50	0.0917	4.654	0.4269
	0.59	7.50	0.0549	7.615	0.418
	0.98	7.50	0.0345	8.387	0.2898
	1.37	7.50	0.0523	8.645	0.4523
	1.76	7.50	0.0367	8.35	0.3070
	2.16	7.50	0.0608	7.549	0.4596
	2.55	7.50	0.0389	7.778	0.3026
	π	7.50	0.119	5.360	0.6419

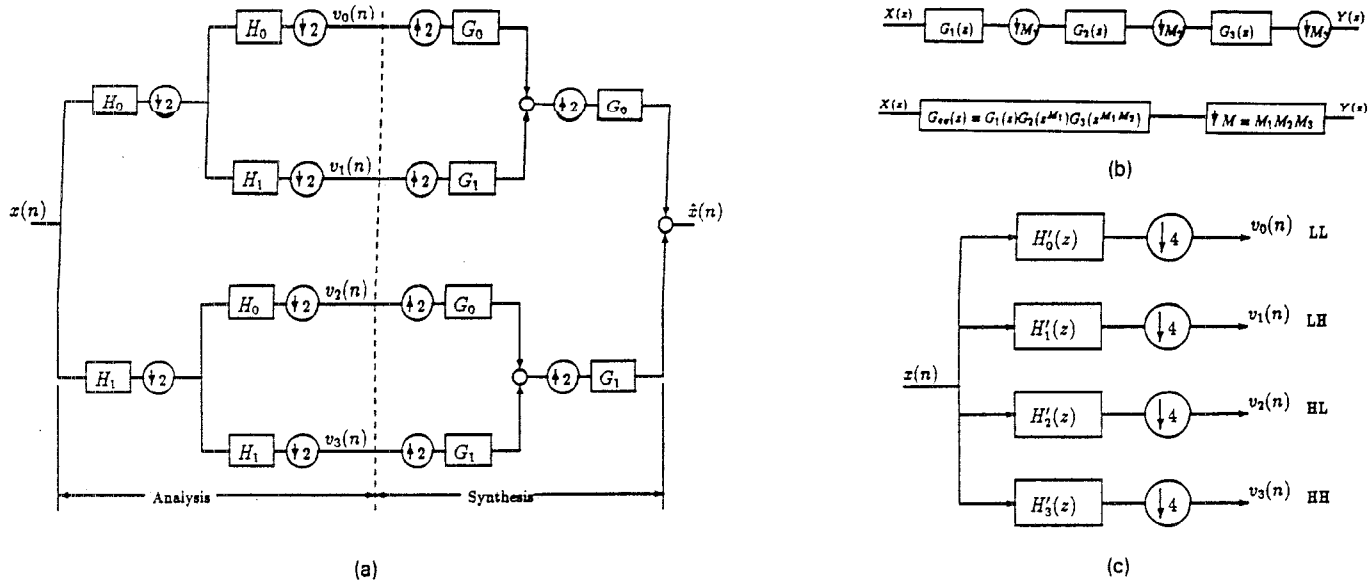


Fig. 9 (a) Four-band hierarchical filter bank, (b) equivalent structures for filters separated by down-samplers, and (c) four-band filter bank equivalent to (a).

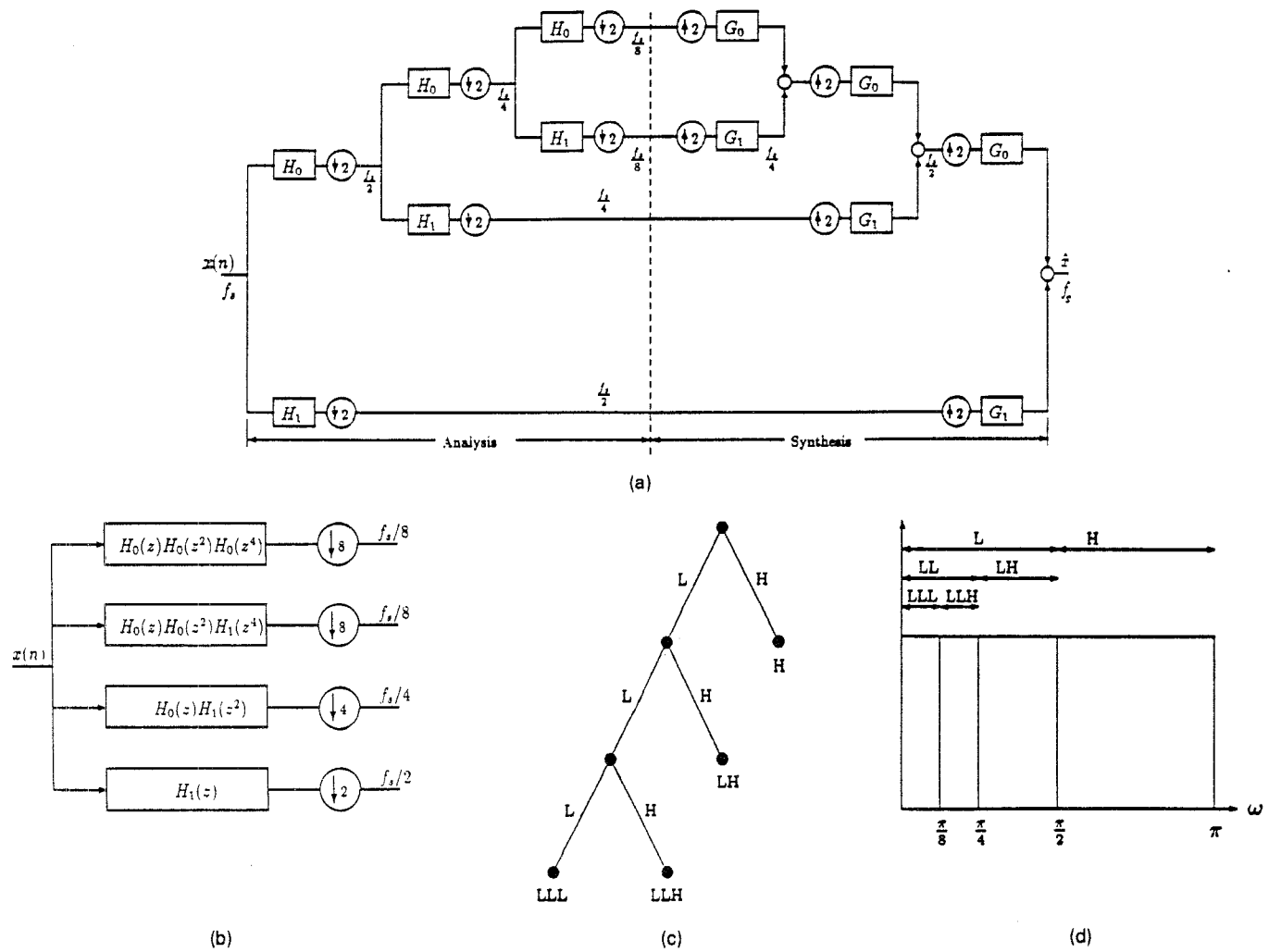


Fig. 10 (a) Four-band dyadic filter bank and (b) its equivalent for the analysis stage as a direct unequal band filter bank, (c) schematic of subband tree, and (d) frequency band split using an ideal two-band PR filter bank in the tree.

4.1 *Representative Two-Band Filter Banks*

The binomial filters introduced earlier provide basis functions for a paraunitary two-band filter bank. The $(N+1)$ tap low-pass filter $H_0(z)$ can be realized as a linear combination of the first $(N+1)/2$ binomial filters,[23]

$$H_0(z)=\sum_{r=0}^{N-1/2}\theta_r X_r(z)=\sum_{r=0}^{N-1/2}\theta_r(1+z^{-1})^{N-r}(1-z^{-1})^r$$
$$=(1+z^{-1})^{N+1/2}F(z)\ . \qquad (42)$$

The imposition of the PR paraunitary requirements of Eq. (40) leads to a set of $(N+1)/2$ nonlinear algebraic equations in the $(N+1)/2$ unknowns, θ_0, θ_1, ..., $\theta_{N-1/2}$. As discussed in Ref. 23, the resulting binomial quadrature mirror filter (QMF) $H_0(z)$ is the unique maximally flat magnitude square solution. Both $H_0(z)$ and $H_1(z)$ can be realized very efficiently using binomial networks with tap weights corresponding to the $(N+1)/2$ coefficients $\{\theta_r\}$. Furthermore, these binomial QMFs are identical to the FIR wavelet filters proposed by Daubechies.[22,25]

For our present purposes, we want to evaluate the time-frequency localization properties of some known filter banks. Table 4 lists these characteristics for three different eight-tap two-band filter banks: the binomial QMF, the Smith and Barnwell conjugate quadrature filter (CQF)[19] and the multiplierless PR QMF.[26] Tables 5 and 6 continue this comparison for hierarchical structure four-band (22-tap product filters) and eight-band (50-tap product filters) configurations. In all these cases, the multiplierless structure has the best time-frequency product $\sigma_n\sigma_\omega$, followed by the Smith and Barnwell CQF and the binomial QMF. As expected, longer duration filters have a narrower σ_ω and a wider σ_n. Again, as expected, the eight-band, eight-tap block transforms (Table 2) have a much narrower σ_n than any of the eight-band tree-structured filter banks, but very poor frequency localization.

Figure 11 displays the impulse responses of the product filters of the two-band binomial QMF-based hierarchical tree for the two-, four-, and eight-band case, and Fig. 12 displays the basis functions of the 2×2, 4×4, and 8×8 DCT. Figure 13 shows the corresponding frequency responses. These demonstrate the drawbacks of blindly repeating a two-band PR-QMF module in a hierarchical subband tree. The time spread increases considerably while the time-frequency product degrades.[8,24,27] This suggests two possibilities: either design the M-band, single-level structure directly or, in using the hierarchical tree structure, monitor the PR-QMF module from level to level.

5 Wavelets and Time-Frequency Decomposition

5.1 *Time-Frequency "Tiling"*

The orthonormal wavelets have recently been advanced as a new mathematical tool for multiresolution decomposition of continuous-time signals with potential applications in several fields.[22,28-30] The wavelet transform is a mapping with superior time-frequency localization compared with the STFT.

The wavelet transform is defined in terms of *dilation* and *translation* of a prototype kernel or *mother function* $\psi(t)$. Its attributes are suggested by the scaling property of Fourier transforms. If $\psi(t)\leftrightarrow\Psi(\Omega)$ are a transform pair, then

Table 4 Time-frequency localizations of three different eight-tap two-band PR-QMF banks.

	$\bar{\omega}$	$\bar{n}$	σ_ω^2	σ_n^2	$\sigma_\omega^2\times\sigma_n^2$
B-QMF (8-tap)	0	1.46	0.9468	0.6025	0.5704
	π	5.54	0.9468	0.6025	0.5704
Multiplierless	0	2.50	0.9743	0.3750	0.3654
(8-tap)	π	4.50	0.9743	0.3750	0.3654
Smith-Barnwell	0	4.17	0.9174	0.5099	0.4678
(8-tap)	π	2.83	0.9174	0.5099	0.4678

Table 5 Time-frequency localizations of hierarchical subband trees for two-level (4 bands).

	$\bar{\omega}$	$\bar{n}$	σ_ω^2	σ_n^2	$\sigma_\omega^2\times\sigma_n^2$
B-QMF Hierarchical	0	4.05	0.2526	2.7261	0.6886
4 Band Tree	1.23	12.88	0.1222	3.8269	0.4676
(22 tap product	1.91	16.28	0.1222	2.7757	0.3392
filters	π	8.80	0.2526	2.2622	0.5714
Multiplierless	0	7.50	0.2747	1.5817	0.4345
(22 tap product	1.24	11.50	0.1346	2.1683	0.2918
filters)	1.90	13.49	0.1346	2.1675	0.2918
	π	9.50	0.2747	1.5818	0.4345
Smith-Barnwell	0	12.45	0.2339	2.1458	0.5019
(22 tap product	1.22	9.88	0.1077	2.9463	0.3173
filters)	1.92	8.45	0.1077	3.0185	0.3251
	π	11.22	0.2339	2.0772	0.4859

Table 6 Time-frequency localizations of hierarchical subband trees for three-level (eight bands) cases.

	$\bar{\omega}$	$\bar{n}$	σ_ω^2	σ_n^2	$\sigma_\omega^2\times\sigma_n^2$
B-QMF Hierarchical	0	9.12	0.0644	11.726	0.7552
8 Band Tree	0.63	26.96	0.0490	15.953	0.7817
(50 tap product	1.01	34.11	0.0961	11.326	1.0884
filters)	1.45	19.65	0.0496	9.7846	0.4853
	1.68	22.56	0.0496	10.510	0.5213
	2.13	37.99	0.0961	12.013	1.1544
	2.52	31.54	0.0490	14.950	0.7326
	π	14.36	0.0644	10.777	0.6940
Multiplierless	0	17.53	0.0724	6.3415	0.4591
(50-tap product	0.64	25.46	0.0688	8.8171	0.6066
filters)	1.02	29.46	0.1193	9.1282	1.0890
	1.45	21.54	0.0558	7.2005	0.4018
	1.68	23.47	0.0558	7.2099	0.4023
	2.11	31.53	0.1193	9.1234	1.0884
	2.50	27.54	0.0688	8.8269	0.6073
	π	19.47	0.0724	6.3371	0.4588
Smith-Barnwell	0	28.86	0.0591	8.6494	0.5112
(50-tap product	0.6137	24.03	0.0321	11.837	0.3800
filters)	0.9951	21.22	0.0688	12.623	0.8685
	1.4488	26.55	0.0436	9.5939	0.4183
	1.6927	25.32	0.0436	9.6769	0.4219
	2.1465	19.57	0.0688	12.599	0.8668
	2.5279	22.51	0.0321	11.912	0.3824
	π	27.93	0.0591	8.5379	0.5046

$$\psi_{ab}(t)=\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)\leftrightarrow\Psi_{ab}(\Omega)$$
$$=\sqrt{a}\Psi(a\Omega)\ \exp(-jb\Omega)\ , \qquad (43)$$

with $a\geq0$, $-\infty<b<\infty$. Thus, a contraction in time is ac-

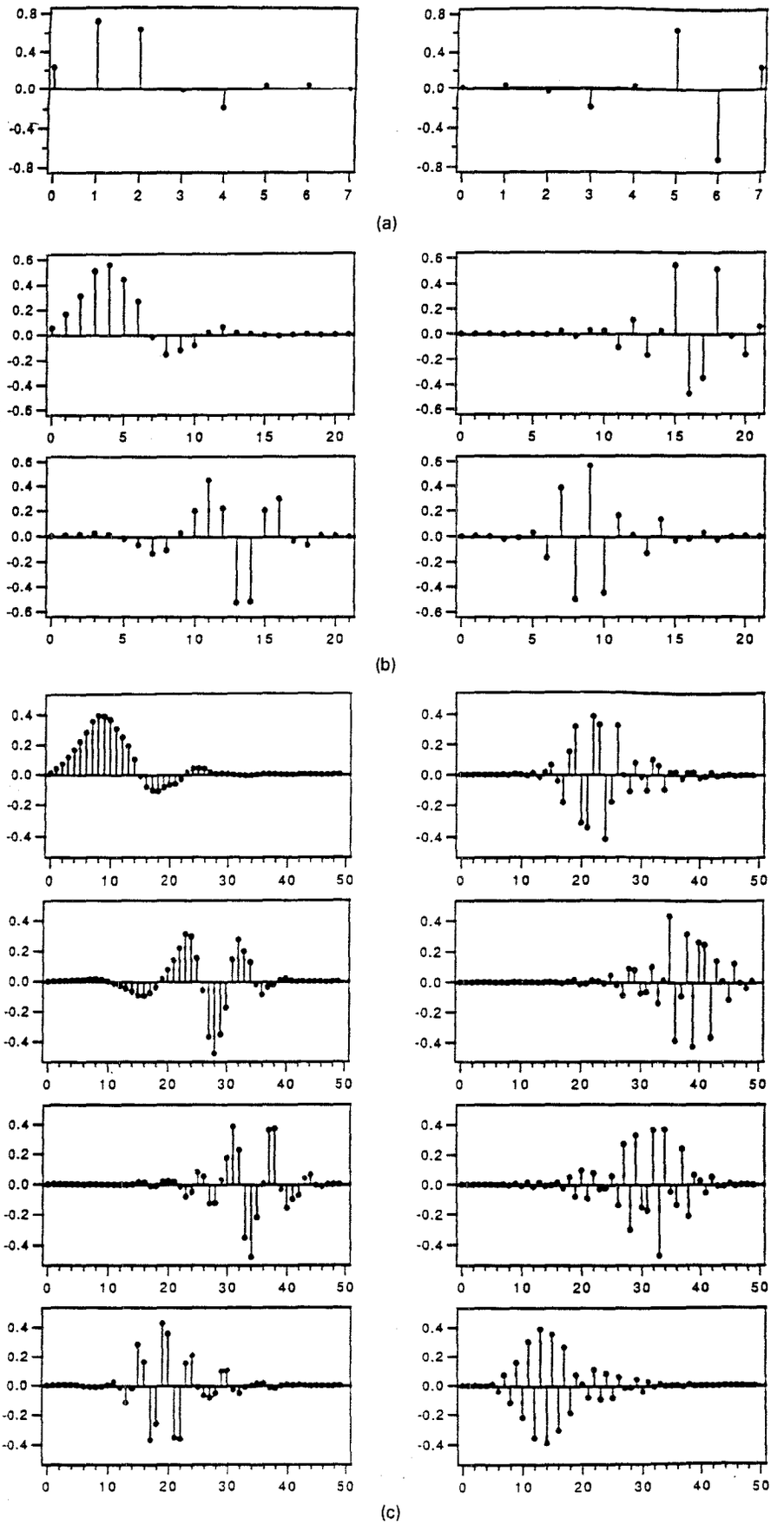


Fig. 11 Impulse responses of the product filters of the two-band binomial QMF-based hierarchical tree for the (a) two-band, (b) four-band, and (c) eight-band case.

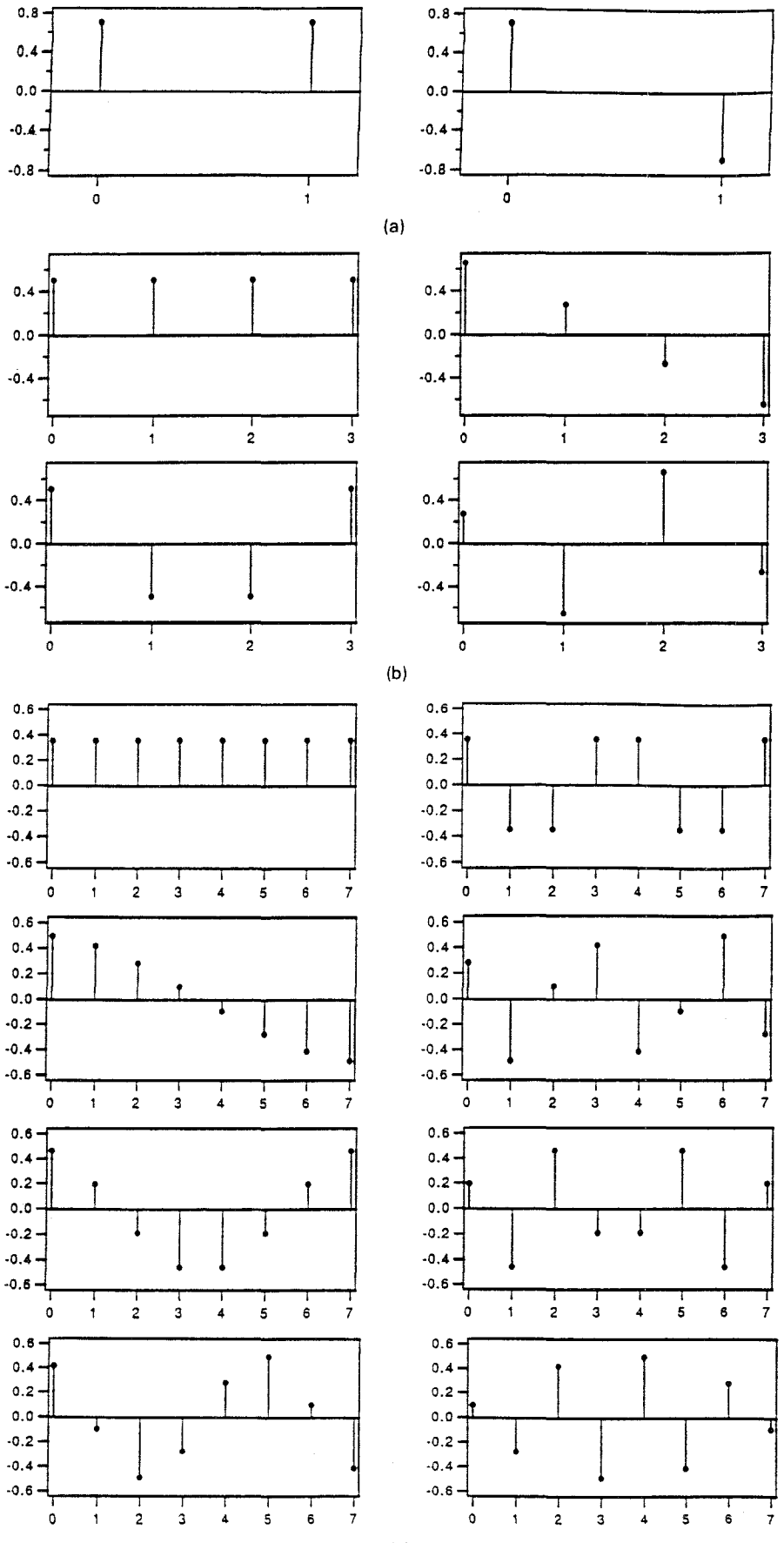


Fig. 12 Basis functions of the 2×2, 4×4, and 8×8 DCT.

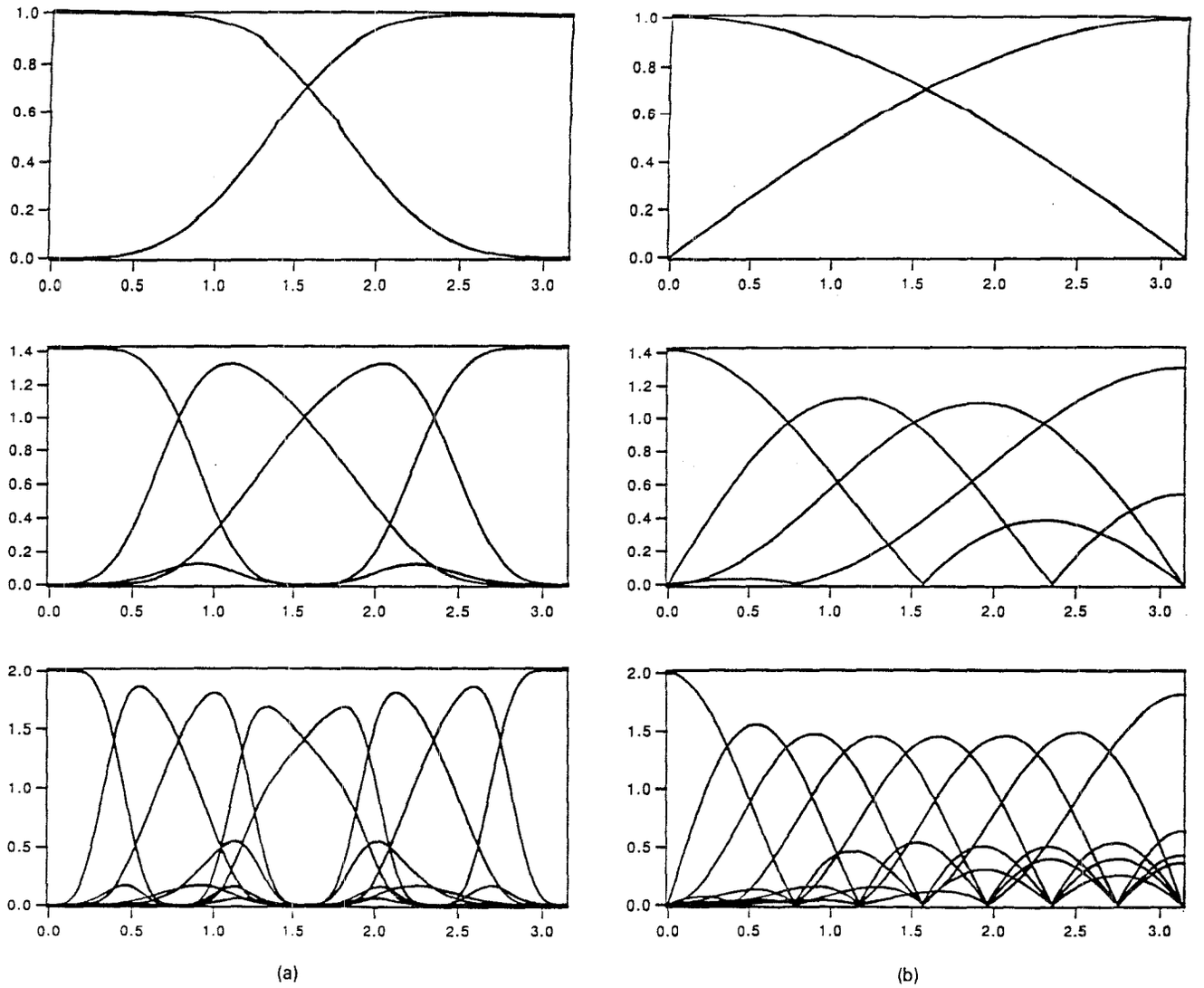


Fig. 13 Frequency responses of (a) functions in Fig. 11 and (b) functions in Fig. 12.

companied by a dilation in frequency and the converse. Here, the scale and shift parameters are a and b, respectively.

The wavelet transform maps a function $f(t)$ on a time-scale space (a,b) by

$$W_f(\mathrm{a,b}) = \int_{-\infty}^{\infty} \psi_{ab}(t) f(t)\ \mathrm{d}t \triangleq \langle \psi_{ab}, f \rangle \ . \tag{44}$$

This transform can be inverted provided $\psi(t)$ satisfies the *admissibility*[22]; then,

$$f(t) = \frac{1}{C} \int_{-\infty}^{\infty} \int_{0}^{\infty} \frac{\mathrm{d}a\ \mathrm{d}b}{a^2} W_f(a,b) \psi_{ab}(t) \ , \tag{45}$$

$$C = \int_0^{\infty} \frac{|\Psi(\Omega)|^2\ \mathrm{d}\Omega}{\Omega} \ . \tag{46}$$

Hence, $\Psi(0)$ must be zero so that $\psi(t)$ behaves as the impulse response of a bandpass filter.

The time-frequency localization of the wavelet transform is distinctly different from that of the STFT. Define $\bar{t}$ and $\bar{\Omega}$ as the center of mass of $\psi(t)$, $\Psi(\Omega)$ on $(-\infty,\infty)$, and $[0,\infty]$, respectively. Then the rms spreads for the prototype are

$$\sigma_{\Omega}^2 = \frac{\dfrac{1}{\pi}\displaystyle\int_0^{\infty} (\Omega - \bar{\Omega})^2 |\Psi(\Omega)|^2\ \mathrm{d}\Omega}{E} \ ,$$

$$\sigma_T^2 = \frac{\displaystyle\int_{-\infty}^{\infty} (t - \bar{t})^2 |\psi(t)|^2\ \mathrm{d}t}{E} \ . \tag{47}$$

This wavelet function is centered at $\bar{t}$ and $\bar{\Omega}$ in the time-frequency plane with spreads σ_T and σ_{Ω}. It follows that $\psi_{ab}(t) \leftrightarrow \Psi_{ab}(\Omega)$ is centered at $(\bar{t}_{ab}, \bar{\Omega}_{ab}) = (a\bar{t} + b, \bar{\Omega}/a)$ with spread

$$\sigma_{\Omega(ab)}^2 = \frac{1}{a^2} \sigma_{\Omega}^2 \ ,$$

$$\sigma_{T(ab)}^2 = a^2 \sigma_t^2 \ , \tag{48}$$

and resolution cell $\sigma_{T(ab)}\sigma_{\Omega(ab)} \geq 1/2$. Thus, the resolution, i.e., localization, in time and frequency depends on this scale parameter a. The shape of the resolution cell depends on a although the cell area is constant. This can be contrasted with the STFT that has a constant shape resolution. These are shown in Fig. 14, and the span of these cells in the time-frequency plane is called a *tiling* of the plane.

An often-quoted example of a wavelet is the second derivative of a Gaussian,

$$\psi(t) = (1-t^2)\ \exp(-t^2/2) \leftrightarrow \Psi(\Omega)$$
$$= \sqrt{2\pi}\Omega^2\ \exp(-\Omega^2/2)\ .$$

This mother function has excellent localization in time and frequency and satisfies the admissibility condition. But it is not of compact support, and is not complete. For this case, we calculate ($\bar{t}=0$, $\sigma_T^2 = 7/6$) and ($\bar{\Omega} = 1.505$, $\sigma_\Omega^2 = 0.23646$), resulting in a cell $\sigma_T\sigma_\Omega = 0.525$.

5.2 *The Discrete Wavelet Transform*

The discretization in the discrete wavelet transform (DWT) refers to the scaling and translation parameters, $a = 2^m$ and $b = n2^m$, which provide an octave grid in the time-frequency plane. This results in the wavelet family,[22,29]

$$\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n)\ . \tag{49}$$

The discrete wavelet transform expansion and its inversion take the form:

$$d_{m,n} = \int_{-\infty}^{\infty} f(t)\psi_{m,n}(t)\ dt \triangleq \langle f(t), \psi_{m,n}(t)\rangle \tag{50}$$

$$f(t) = \sum_{m=0}^{\infty}\sum_{n=-\infty}^{\infty} d_{m,n}\psi_{m,n}(t)\ . \tag{51}$$

Two main connections exist between orthonormal wavelets of compact support and paraunitary dyadic subband tree structures. The first deals with the construction of the wavelet function $\psi(t)$ and the second with the calculation of the wavelet coefficients $\{d_{m,n}\}$ for a given function $f(t)$ and given wavelet family. These results are derived in the literature.[22,13] We only summarize the main results.

5.3 *Construction of Wavelets from Unitary PR QMFs*

The construction of wavelets from unitary PR QMFs requires the following:

1. The wavelets are orthonormal in both indices. They are orthonormal in time n at the same scale m, and orthonormal across scales (intra- and interscale orthonormalities),

$$\langle \psi_{m,n}(t), \psi_{m',n'}(t)\rangle = \delta_{m-m'}\delta_{n-n'}\ . \tag{52}$$

2. The complementary scaling function $\phi(t) \leftrightarrow \Phi(\Omega)$ is orthonormal within the same scale (intrascale orthonormality only),

$$\langle \phi_{mn}(t), \phi_{mn'}(t)\rangle = \delta_{n-n'}\ , \tag{53}$$

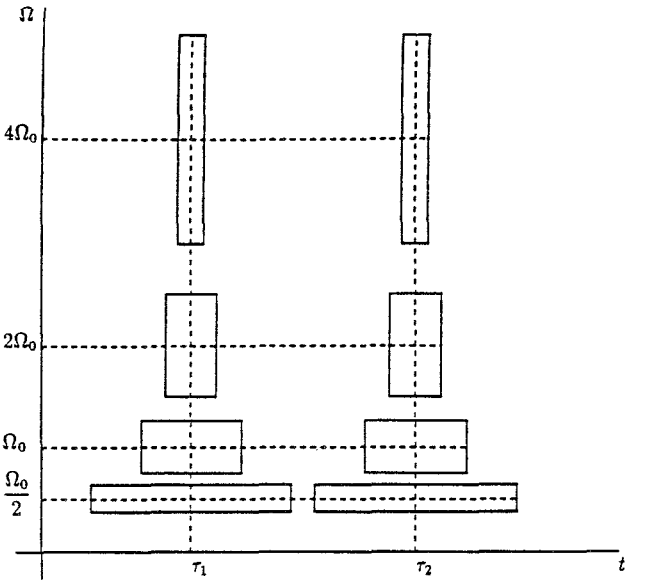


Fig. 14 Time-frequency plane showing resolution cells for wavelet transform.

where

$$\phi_{mn}(t) = 2^{-m/2}\phi(2^{-m}t - n)\ . \tag{54}$$

The scaling function is a low-pass function.

3. Complementary property of the wavelet and scaling bases,

$$\langle \psi_{mn}(t), \phi_{m'n'}(t)\rangle = 0\ . \tag{55}$$

4. Let $h_0(n)$ and $h_1(n)$ constitute a two-band paraunitary filter bank, Eqs. (40) and (41), with the added property that $H_0[\exp(j\omega)] = 0$ at $\omega = \pi$ (i.e., $H_0(z)$ has at least one zero at $z = -1$). Then it can be shown that $\phi(t)$ and $\psi(t)$ can be constructed from the discrete-time filters or the interscale coefficients, via the fundamental wavelet equations of the containment property,

$$\phi(t) = \sum_n h_0(n)\phi(2t-n) \leftrightarrow \Phi(\Omega)$$
$$= \prod_{k=1}^{\infty} H_0[\exp(j\omega/2^k)]\ ,$$
$$\psi(t) = \sum_n h_1(n)\phi(2t-n) \leftrightarrow \Psi(\Omega)$$
$$= H_1[\exp(j\omega/2)]\prod_{k=2}^{\infty} H_0[\exp(j\omega/2^k)]\ . \tag{56}$$

It turns out that the orthonormality and finite support of the scaling and wavelet functions is set up by the orthonormality and finite duration of $h_0(n)$ and $h_1(n)$. Equation (56) provides the basis for the construction of wavelet families. One simply starts with any paraunitary, compact support (i.e., finite duration) filters. The Fourier transforms of the wavelet and scaling functions are then obtained by the infinite products in Eq. (56)!

5.4 *Wavelet Expansion and the Dyadic Subband Tree*

The main result here is that we can use a paraunitary dyadic subband tree to compute the coefficients $\{d_{m,n}\}$ in the wavelet expansion of a continuous-time function.

The basic construction is as follows: Suppose $f(t)$ can be represented by a sum of orthonormal scaling functions at full resolution, i.e.,

$$f(t)=\sum_n c_{0,n}\phi(t-n)=\sum c_{0,n}\phi_{0n}(t) \ , \tag{57}$$

where the scaling coefficients

$$c_{0,n}=\langle f(t),\phi_{0n}(t)\rangle \ . \tag{58}$$

This is the starting point for the wavelet decomposition. Now suppose that $c_{0,n}$ is the input to the dyadic subband tree shown in Fig. 10, and $h_0(n)$ and $h_1(n)$ are paraunitary filters that generated $\Phi(\Omega)$ and $\Psi(\Omega)$ in Eq. (56).

By the fundamental wavelet equation, $f(t)$ can be decomposed into a (coarse) lower resolution scaling expansion and a lower resolution wavelet expansion that carries the detail

$$f(t)=\sum_n d_{1,n}\psi_{1n}(t)+\sum_n c_{1,n}\phi_{1n}(t) \ .$$

The coarse approximation at the half resolution can in turn be decomposed into yet a coarser approximation at the quarter resolution and a detail (wavelet), or

$$f(t)=\sum_n d_{1,n}\psi_{1n}(t)+\sum_n d_{2,n}\psi_{2n}(t)+\sum_n c_{2,n}\phi_{2n}(t) \ .$$

Thus, $f(t)$ can be represented as a low-pass approximation at the $1/L$ resolution plus a sum of L detail wavelet components at successively finer resolutions. This decomposition can be continued indefinitely. As shown in Fig. 10, the coefficients in this pyramid expansion are simply the output of the paraunitary dyadic subband tree. Hence, the terminology *fast wavelet transform.*

The fly in the ointment here is the initialization or the feeding of the subband tree by $\{c_{0,n}\}$. If this starting point, Eq. (58) is only an approximation, then the expansion that follows is itself an approximation. As a case in point, suppose $f(t)$ is a band-limited signal. Then

$$\phi_{0n}(t)=\frac{\sin\pi(t-n)}{\pi(t-n)}$$

is an orthonormal basis, and

$$f(t)=\sum f(n)\frac{\sin\pi(t-n)}{\pi(t-n)} \ .$$

In this case, $h_0(n)$ and $h_1(n)$ must be ideal low-pass and high-pass filters. If $f(n)$ is the input to the dyadic subband tree with filters that only approximate the ideal filters, then the resulting coefficients or subband signals $\{d_{m,n}\}$ and $\{c_{m,n}\}$ are themselves only approximations to the exact values.

It is hard to justify performing a wavelet transform on sampled data. The discrete-time subband decomposition is the natural technique for these types of signals.

Table 7 Time-frequency localizations of six-tap wavelet filters and corresponding scaling and wavelet functions.

		Daubechies *	Mostregular †	Coiflet ‡
Scaling Function	σ_T^2	0.134	0.143	0.086
	σ_Ω^2	5.22	5.77	11.86
	$\sigma_T^2\sigma_\Omega^2$	0.699	0.825	1.02
Wavelet Function	σ_T^2	0.178	0.188	0.108
	σ_Ω^2	8.97	11.70	39.36
	$\sigma_T^2\sigma_\Omega^2$	1.596	2.199	4.25
Low-Pass PR-QMF	σ_n^2	0.453	0.470	0.305
	σ_ω^2	0.987	0.996	1.059
High-Pass PR-QMF	σ_n^2	0.453	0.470	0.305
	σ_ω^2	0.987	0.996	1.059

* See Ref. 22 and 23.

† See Ref. 33.

‡ See Ref. 32.

5.5 *Time-Frequency Resolution for Wavelet Families*

We have seen that orthonormal wavelet families can be generated by any pair of two-band paraunitary filters $H_0(z)$ and $H_1(z)$, with $H_0(-1)=0$. The generation of Daubechies wavelets are given in Ref. 22, and it shows that these wavelet filters are identical to the binomial QMF. These filters have the maximally flat magnitude square responses.[23,31] In Refs. 32 and 33, other wavelet families (e.g., the most regular, Coiflets) are devised by imposing other requirements on $H_0(z)$.

Table 7 compares the time-frequency resolutions of scaling and wavelet functions for three wavelet families generated by six-tap paraunitary filters—the Daubechies, most regular, and Coiflet—along with the localization properties of the progenitor discrete-time filters. As seen in Eq. (56), the Fourier domain equalities require infinite product terms theoretically. The localization measures in this table assumed the maximum product order of $k=8$ in Eq. (56). It considered the range of $-2^k\pi \leq \Omega \leq 2^k\pi$. Table 7 demonstrates that the time-frequency localizations are important measures in the evaluation of a wavelet family as an analog filter bank. In particular, the role of regularity in wavelet transforms should be evaluated for signal processing applications.

6 Discussion and Conclusions

We examined the discrete-time uncertainty principle, its measure by a resolution cell, and the lower bounds for different classes of signals. We also evaluated the time-frequency resolutions of some known orthogonal signal decomposition techniques: block transforms, lapped orthogonal transforms, subband filter banks, and wavelets.

Historically, the design of transform bases and filter banks has emphasized either the time or frequency domain with orthogonality as their main structure. It is well observed and understood in visual signal processing and coding applications that the behavior of the transform basis or filters should be monitored jointly in the time and frequency domains. It is expected that this point will be considered in the future designs.

The FIR two-band orthonormal filter banks have a vital role as the interscale coefficients in the design of compactly supported orthonormal wavelet transform bases. The wavelet theory emphasizes the differentiability or regularity of the

basis functions in the design. Wavelet regularity implies a flat frequency response for the wavelet filters at $\omega=0$ and $\omega=\pi$. In Table 7, the time and frequency localizations of wavelet filters along with the corresponding wavelet and scaling functions were evaluated for three different families proposed by Daubechies. In addition to the regularity measure, new performance measures may be needed in the design of wavelet bases for signal processing applications.

The transform bases or filter banks consist of a set of functions. The interband leakage or overlapping of frequency functions should also be considered in the design. This means a good localization at the desired region of the time-frequency plane for all the functions in the basis.

We conclude that the overlapping transform basis or filter bank design has significantly more degrees of freedom than has been utilized. We expect more flexible basis designs in the future for more efficient spectral decomposition of signal sources.

7 Appendix

7.1 *Discrete-Time Uncertainty*[6,7]

Let $f(n) \leftrightarrow F[\exp(j\omega)]$ be a discrete-time Fourier transform pair, with $f(n)$ real. For convenience let $\bar{n}=0$, and note that

$$nf(n) \leftrightarrow j\frac{dF}{d\omega} , \tag{59}$$

so that

$$\sum_n n^2|f(n)|^2=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{dF}{d\omega}\right|^2 d\omega . \tag{60}$$

By the Schwartz' inequality, we have

$$|I|^2=\left|\frac{1}{2\pi}\int_{-\pi}^{\pi}\omega F^*[\exp(j\omega)]\frac{dF}{d\omega}d\omega\right|^2$$

$$\leq\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\omega^2\left|F[\exp(j\omega)]\right|^2 d\omega\right\}$$

$$\times\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{dF}{d\omega}\right|^2 d\omega\right) . \tag{61}$$

Integrating by parts,

$$I=\frac{1}{2\pi}\{\omega|F[\exp(j\omega)]|^2\}_{-\pi}^{\pi}$$

$$-\frac{1}{2\pi}\int_{-\pi}^{\pi}F[\exp(j\omega)](\omega\, dF^*+F^*\, d\omega)$$

$$=|F(-1)|^2-\frac{1}{2\pi}\int_{-\pi}^{\pi}\omega F[\exp(j\omega)]\frac{dF^*}{d\omega}$$

$$-\frac{1}{2\pi}\int_{-\pi}^{\pi}|F[\exp(j\omega)]|^2 d\omega$$

$$=|F(-1)|^2-I^*-E \tag{62}$$

or

$$I+I^*=|F(-1)|^2-E .$$

But,

$$|I+I^*|\leq 2I$$

for any complex number, therefore,

$$\left|\frac{I}{E}\right|=\sigma_\omega\sigma_n\geq\frac{1}{2E}|E-|F(-1)|^2|=\frac{1}{2}|1-\mu| .$$

7.2 *Calculation of σ_ω for Gaussian*

The frequency localization of Gaussian function can be found as:

$$\sigma_\omega^2=\frac{\frac{1}{2\pi}\int_{\pi}^{\pi}\omega^2|F[\exp(j\omega)]|^2 d\omega}{E}$$

$$=\frac{K}{\pi}\int_0^{\pi}\omega^2\exp(-\omega^2/2\sigma^2)\, d\omega . \tag{63}$$

Let $x=\omega/\sigma$ and integrate them by parts to obtain:

$$\sigma_\omega^2=\frac{-K\sigma^3}{\pi}\left\{[-x\exp(-x^2/2)]_0^{\pi/\sigma}+\int_0^{\pi/\sigma}\exp(-x^2/2)\, dx\right\}$$

$$=\frac{K\sigma^3}{\pi}\sqrt{2\pi}\,\mathrm{erf}(\pi/\sigma)-\frac{K\sigma^3}{\pi}(\pi/\sigma)\exp(\pi^2/2\sigma^2)$$

$$=\sigma^2[1-K\exp(-\pi^2/2\sigma^2)]=\sigma^2(1-\mu)$$

$$\sigma_\omega=\sigma\sqrt{1-\mu} . \tag{64}$$

7.3 *Derivation of Eqs. (25) and (27) for Binomial*

The derivation of Eqs. 25 and 27 for the binomial:

$$E=\sum_{k=0}^{N}\binom{N}{k}=\frac{1}{2\pi}\int_{-\pi}^{\pi}|F[\exp(j\omega)]|^2 d\omega$$

$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}(2^{2N})\left(\cos\frac{\omega}{2}\right)^{2N} d\omega$$

$$=\frac{2}{\pi}(2)^{2N}\int_0^{\pi/2}(\cos x)^{2N}\, dx . \tag{65}$$

From the Chemical Rubber Company Tables,

$$\int_0^{\pi/2}(\cos x)^{2N}\, dx = \left[\frac{1.3.5...(2N-1)}{2.4.6...(2N)}\right]\left(\frac{\pi}{2}\right)$$

and Eq. (25) results. Then

$$\sigma_n^2=\frac{\sum(n-\bar{n})^2|f(n)|^2}{\sum|f(n)|^2}=\frac{\sum n^2|f(n)|^2}{E}-(\bar{n})^2=\frac{B}{E}-(N/2)^2 , \tag{66}$$

where

$$B=\sum n^2|f(n)|^2=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{dF}{d\omega}\right|^2 d\omega$$

$$=\frac{(2)^{2N-1}N^2}{\pi}\int_0^{\pi}(\cos x)^{2(N-1)}\,dx$$

$$=(2)^{2N}\frac{N^2}{4}\left[\frac{2.4...2N}{1.3.5...(2N-1)}\right]. \quad (67)$$

This gives

$$\frac{B}{E}=(N/2)^2[2N/(2N-1)] \quad (68)$$

Substituting this last result into Eq. (25) gives

$$\sigma_n^2=(N/2)^2\left(\frac{2N}{2N-1}-1\right)=(N/2)^2\frac{1}{(2N-1)}. \quad (69)$$

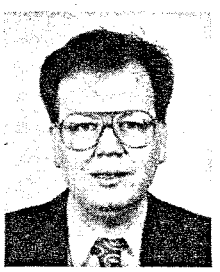
Acknowledgment

The authors would like to thank Mr. M. V. Tazebay for providing some of the figures.

References

1. H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty-II," *Bell Sys. Tech. J.*, pp. 65–84 (Jan. 1961).
2. A. Papoulis, *Signal Analysis,* McGraw-Hill, New York (1977).
3. L. Cohen, "Time-frequency distribution—a review," *Proc. IEEE* **77,** 941–981 (1989).
4. D. Gabor, "Theory of communication," *J. IEE* **93,** 429–457 (1946).
5. R. A. Haddad and T. W. Parsons, *Digital Signal Processing: Theory, Applications, and Hardware,* Computer Science Press, New York (1991).
6. Y. Liu, H. Caglar, and A. N. Akansu, *Time-Frequency Localization and Optimal PR-QMF Design* (1992).
7. L. C. Calvez and P. Vilbe, "On the uncertainty principle in discrete signals," *IEEE Trans. Circ. Syst.-II: Analog and Digital Signal Processing* **39**(6), 394–395 (June 1992).
8. Y. Liu and A. N. Akansu, "An evaluation of time-frequency localization in transforms and filter banks," *Proc. IEEE ICASSP* **III,** pp. 261–264 (1993).
9. R. A. Haddad, "A class of orthogonal nonrecursive binomial filters," *IEEE Trans. Audio Electroacoust.* **AU-19**(4), 296–304 (Dec. 1971).
10. P. P. Vaidyanathan, "Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M, having the perfect reconstruction property," *IEEE Trans. Acoust. Speech Sig. Proc.* **ASSP-35,** 476–492 (April 1987).
11. M. Vetterli and D. LeGall, "Perfect reconstruction FIR filter banks: some properties and factorizations," *IEEE Trans. Acoust. Speech Sig. Proc.* **ASSP-37,** 1057–1071 (July 1989).
12. K. Nayebi, T. P. Barnwell III, and M. J. T. Smith, "Time domain filter bank analysis: a new design theory," *IEEE Trans. Sig. Process.* **40**(6), 1412–1429 (June 1992).
13. A. N. Akansu and R. A. Haddad, *Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets,* Academic Press, Inc., New York (1992).
14. H. S. Malvar, *Signal Processing with Lapped Transforms,* Artech House, Boston (1992).
15. P. P. Vaidyanathan, *Multirate Systems and Filterbanks,* Prentice-Hall, New Jersey (1993).
16. H. S. Malvar and D. H. Staelin, "The LOT: transform coding without blocking effects," *IEEE Trans.* **ASSP-37**(4), 553–559 (April 1989).
17. H. S. Malvar, "The LOT: a link between block transform coding and multirate filter banks," *Proc. IEEE ISCAS,* pp. 781–784 (1988).
18. A. N. Akansu and F. E. Wadas, "On lapped orthogonal transforms," *IEEE Trans. Sig. Process.* **40**(2), 439–443 (Feb. 1992).
19. M. J. T. Smith and T. P. Barnwell, "Exact reconstruction techniques for tree-structured subband coders," *IEEE Trans. Acoust. Speech Sig. Proc.* **ASSP-34,** 434–441 (1986).
20. F. Mintzer, "Filters for distortion-free two-band multirate filter banks," *IEEE Trans. ASSP* **33,** 626–630 (June 1985).
21. H. Caglar, Y. Liu, and A. N. Akansu, "Statistically optimized PR-QMF design," *Proc. SPIE* **1605,** 86–94 (1991).
22. I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math.* **41,** 909–996 (1988).
23. A. N. Akansu, R. A. Haddad, and H. Caglar, "The binomial QMF-wavelet transform for multiresolution signal decomposition," *IEEE Trans. Sig. Process.* **41,** 13–19 (Jan. 1993).
24. A. N. Akansu and Y. Liu, "On signal decomposition techniques," *Opt. Eng.* **30**(7), 912–920 (July 1991).
25. R. A. Ansari, C. Guillemot, and J. F. Kaiser, "Wavelet construction using Lagrange halfband filters," *IEEE Trans. Circ. Syst.* **CAS-38,** 1116–1118 (Sept. 1991).
26. A. N. Akansu, "Multiplierless suboptimal PR-QMF design," *Proc. SPIE* **1818,** 723–734 (1992).
27. R. Coifman and Y. Meyer, *Orthonormal Wave Packet Bases,* Dept. of Math., Yale Univ., CT, preprint.
28. O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Sig. Process. Mag.* **8,** 14–38 (Oct. 1991).
29. S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Patt. Anal. Mach. Intell.* **11,** 674–693 (July 1989).
30. J. M. Combes, A. Grossmann, and P. H. Tchamitchian, Eds., *Wavelets, Time-Frequency Methods and Phase Space,* Springer-Verlag, Berlin (1989).
31. O. Herrmann, "On the approximation problem in nonrecursive digital filter design," *IEEE Trans. Circ. Theory* **CT-18**(3), 411–413 (May 1971).
32. I. Daubechies, "Orthonormal bases of compactly supported wavelets II. variations on a theme,"Tech. Memo. #11217-891116-17, AT&T Bell Labs., Murray Hill, New Jersey (1989).
33. I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inf. Theo.* **36,** pp. 961–1005 (Sept. 1990).

Richard A. Haddad received the BEE, MEE, and PhD degrees in 1956, 1958, and 1962, respectively, from the Polytechnic Institute of Brooklyn. He has been on the faculty of the Polytechnic University since 1961. During his tenure there, he has served in various capacities: From 1981 to 1987 he was the director of the Westchester Graduate Center and presently, he is assistant head of the School of Electrical Engineering and Computer Science. During leaves of absence, he has served as a member of the technical staff at Bell Telephone Laboratories, Whippany, New Jersey, and as first director of the Engineering Division at the Institute National d'Electricite et d'Electronique, Boumerdes, Algeria. He has performed research in adaptive control systems, feedback and sensitivity, and optimal estimation. He is now engaged in teaching and research in one- and multidimensional signal processing. He serves as a consultant to industry and government agencies in the fields of guidance and real-time digital data filtering and estimation. Dr. Haddad is a senior member of IEEE, Eta Kappa Nu, Tau Beta Pi, and Sigma Xi. He is the coauthor of two texts, one on digital signal processing and the other on multiresolution signal decomposition.



Ali N. Akansu received the BS degree from the Technical University of Istanbul, Turkey, in 1980, the MS degree from the Polytechnic Institute New York, Brooklyn, in 1983, and the PhD degree from the Polytechnic University, Brooklyn, New York, in 1987, all in electrical engineering. Since 1987, he has been at the New Jersey Institute of Technology, where he is currently an associate professor of electrical engineering. He was an academic visitor at IBM T. J. Watson Research Center and at GEC-Marconi Electronic Systems Corp. during the summers of 1989 and 1992, respectively. He serves as a consultant to the industry. His current research interests are signal processing, image-video compression, and pattern recognition. Dr. Akansu is an associate editor of *IEEE Transactions on Signal Processing,* a member of the DSP Technical Committee of the IEEE Signal Processing Society, and he is a member of IEEE and Sigma Xi. He is the coauthor of a book on multiresolution signal decomposition.