MIT Media Laboratory Vision and Modeling Technical Report #161.
Appears in: IEEE Trans. Information Theory, vol. 38(2), pp. 587-607, March 1992.

Shiftable Multi-scale Transforms

Eero P. Simoncelli[†]
William T. Freeman[‡]
Edward H. Adelson[§]
David J. Heeger[*]

[†‡§]The Media Laboratory
[†]Dept. of Electrical Engineering and Computer Science
[§]Dept. of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

[*]NASA-Ames Research Center, Moffett Field, CA 94035
and Stanford Psychology Dept, Stanford, CA 94305

February, 1991
(revised July and October, 1991)

Abstract

Orthogonal wavelet transforms have recently become a popular representation for multiscale signal and image analysis. One of the major drawbacks of these representations is their lack of translation invariance: the content of wavelet subbands is unstable under translations of the input signal. Wavelet transforms are also unstable with respect to dilations of the input signal, and in two dimensions, rotations of the input signal. We formalize these problems by defining a type of translation invariance that we call "shiftability". In the spatial domain, shiftability corresponds to a lack of aliasing; thus, the conditions under which the property holds are specified by the sampling theorem. Shiftability may also be considered in the context of other domains, particularly orientation and scale. We explore "jointly shiftable" transforms that are simultaneously shiftable in more than one domain. Two examples of jointly shiftable transforms are designed and implemented: a one-dimensional transform that is jointly shiftable in position and scale, and a two-dimensional transform that is jointly shiftable in position and orientation. We demonstrate the usefulness of these image representations for scale-space analysis, stereo disparity measurement, and image enhancement.

Keywords: wavelet, multi-scale, pyramid, aliasing, sampling, interpolation, orientation, steerable filters, image representation, image processing.

• Reformatted with reduced line spacing, February 1996.

• Inquiries should be addressed to: Dr. Eero Simoncelli, GRASP Laboratory - rm 335C, 3401 Walnut St., Philadelphia, PA 19103. *email:* eero@central.cis.upenn.edu.

1 Introduction

In many signal processing applications, a signal is decomposed into a set of subbands, and the information within each subband is processed more or less independently of that in the other subbands. In classical signal processing, the subbands are each sampled at a rate above the Nyquist limit in order to avoid aliasing. If the subband spectra overlap, then the transform will be oversampled, and thus overcomplete.

Recently it has become popular to use discrete subband decompositions that are critically sampled, i.e., in which the number of samples in the representation is equal to the number in the signal. Quadrature mirror filters (QMF's) and wavelets, which are closely related, allow one to violate the Nyquist criterion without discarding information; indeed these subband decompositions constitute orthogonal subband transforms. They achieve this feat by ensuring that the aliasing errors from all of the subbands cancel when the bands are recombined.

These transforms have proven to be highly successful for efficient image coding [45, 50, 17, 3, 43, 2, 49] and they have been proposed as useful frameworks for many other signal processing tasks. One of the appealing properties of wavelet transforms is that they represent the signal with a set of basis functions that are related by translation and dilation. These operators correspond to common "physical" transformations of signals. Since the basis functions have this simple relationship, one might hope that the transform coefficients would behave in a simple manner when the input signal is translated or dilated. However, the situation turns out to be rather complex.

To illustrate this point, we consider the case of spatial translation. We take as our transform a three-level pyramid constructed using the discrete dyadic Daubechies wavelet kernel [9] of length four. The input signal is shown at the top of figure 1(a). For illustrative purposes, we have chosen this signal to be equal to one of the wavelet basis functions in the second subband. Thus, its transform is a single impulse within that subband. All of the other coefficients in the transform are zero. The coefficients of the three subbands are plotted in figures 1(b)–(d). Now suppose we translate the input signal one sample to the right, as shown in figure 1(e). The coefficients of the transform, plotted in figures 1(f)–(h), change dramatically. Coefficient power is now spread broadly within *and across* the subbands. Thus, the representation is highly dependent on the relative alignment of the input signal with the subsampling lattices. This is hardly the behavior that one would like to see when attempting to do translation-invariant signal processing.

The fact that the transform contents jump around in our example leads us to ask what we can hope to achieve in the way of translation invariance. We cannot literally expect translation invariance in a system based on convolution and subsampling: translation of the input signal cannot produce simple translations of the transform coefficients, unless the translation is a multiple of each of the subsampling factors in the system. It is, however, possible to achieve a weaker form of translation invariance in the sense that all the information represented within a subband remains in that subband as the input signal is translated. A necessary and sufficient condition for this is just the Nyquist criterion. Since critically sampled subband transforms (such as wavelet transforms) typically violate the Nyquist criterion, the information moves from one band to another under translation.

These issues have been raised elsewhere. Strang mentions the translation invariance problem as one of the primary weaknesses of the wavelet transform [42]. In the context of the

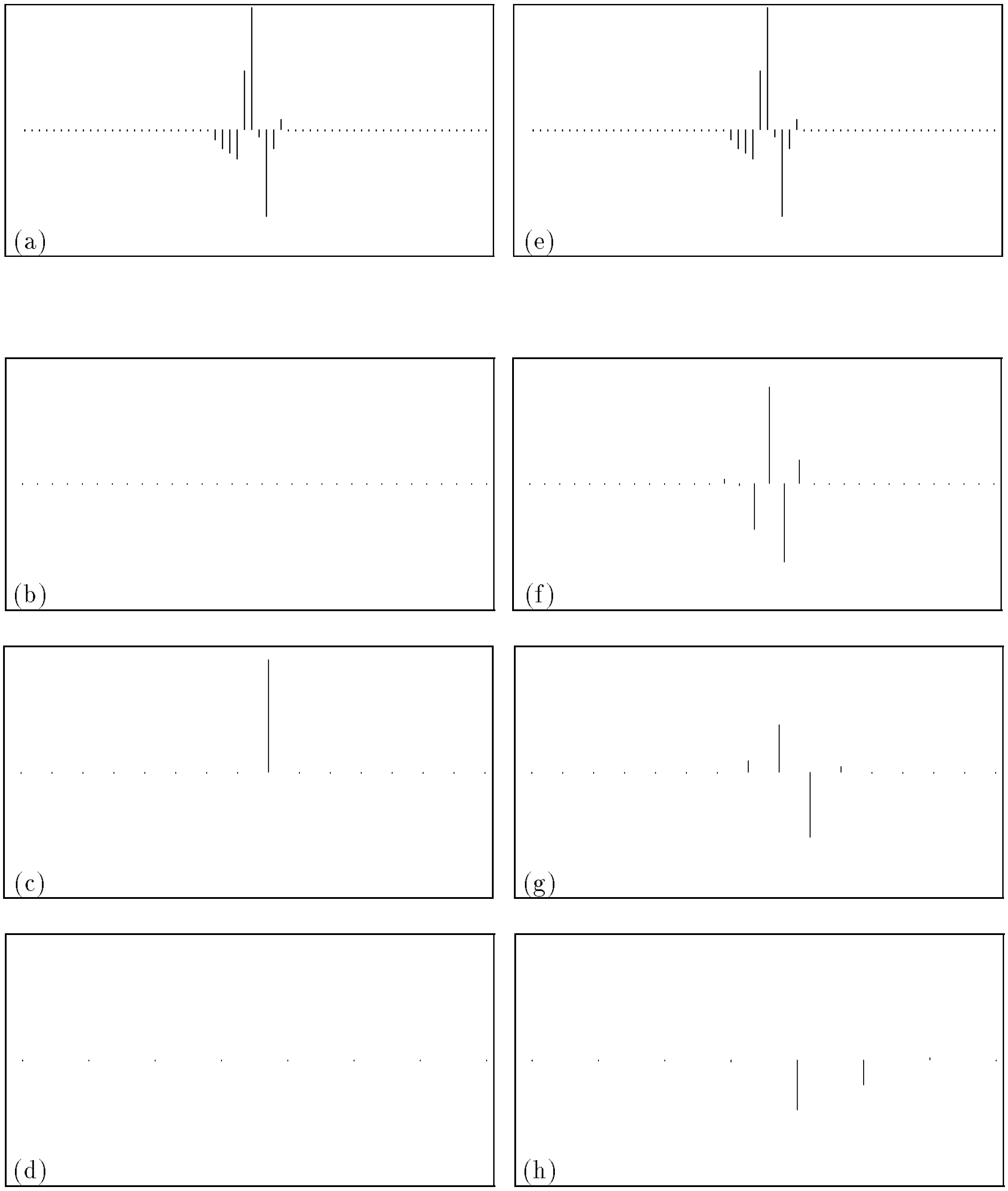


Figure 1: Effect of translation on the wavelet representation of a signal. (a) An input signal, which is equal to one of the wavelet basis functions. (b)–(d) Decomposition of the signal into three wavelet subbands. Plotted are the coefficients of each subband. Dots correspond to zero-value coefficients. (e) The same input signal, translated one sample to the right. (f)–(h) Decomposition of the shifted signal into three wavelet subbands. Note the drastic change in the coefficients of the transform, both within and between subbands.

stereo matching problem, Mallat addressed this problem by building a translation-invariant representation from the zero-crossings of the full-density wavelet subbands [30].

The translation invariance issue serves as a starting point for our discussion, but one of the purposes of this paper is to show that this problem may also be considered in other domains. In two dimensions, wavelet transforms may be computed using basis functions at a particular set of positions, scales and orientations. The same problems of invariance can be stated for the parameters of scale and orientation. The sampling of these domains is analogous to the sampling in the spatial domain, and as with translations, two-dimensional wavelet transforms typically do not behave well under rotations or dilations of the input signal.

The desire for invariance with respect to these simple operations leads us to re-examine the critical sampling constraint of the wavelet representation. More specifically, it may seem that the use of ideal (sinc) kernels could solve the translation invariance problem illustrated in figure 1. But we will show that a critically-sampled transform based on these kernels lacks scale-invariance: the transform does not behave well under dilations of the input signal.

In this paper, we formalize this line of thought by defining the property of "shiftability". We first discuss this property with respect to individual parameters: spatial position, orientation, and scale. We then discuss transformations that are simultaneously shiftable with respect to subsets of these parameters. In order to achieve shiftability, we show that we must relax the critical sampling condition of the wavelet transforms, thus abandoning the elegance of orthogonality. Finally, we implement two example transforms to illustrate these properties and we apply them to several signal and image processing problems.

2 Background

In this section, we define some terminology and review literature on linear transforms and representations related to our topic.

A complete linear transform represents a signal as a weighted sum of *basis* functions. That is, a signal, $f(x)$, is represented as a sum over an indexed collection of functions, $g_i(x)$:

$$f(x) = \sum_i y_i g_i(x), \tag{1}$$

where the y_i are the transform coefficients. These coefficients are computed from the signal by projecting onto a set of functions that we call *projection* functions, $h_i(x)$:

$$y_i = \int_0^{2\pi} dx \;\; h_i(x) f(x). \tag{2}$$

Equation (2) is used to compute the transform of the signal. The original signal may be recovered from the coefficients, y_i, and the basis functions, $g_i(x)$, using equation (1).

In systems based on convolution operations, the projection functions are shifted copies of the (reversed) convolution kernels. From the definition of convolution, the projection functions corresponding to a single convolution and sampling stage are:

$$h_i(x) = h(i\Delta_x - x).$$

In this case, the subscript i indexes the spatial position of its associated basis function, and Δ_x is the sample spacing.

The basis functions, $h_i(x)$, are said to be *linearly independent* if there is no linear combination of them that is zero for all x. In discrete systems, a complete transform with linearly independent basis functions is called *critically sampled*: the coefficient output rate is equal to the input signal sample rate. If the set of basis functions is not linearly independent, then the transform is said to be *overcomplete*. In this case, the corresponding projection functions are not unique. We will refer to a transform for which $h_i(x) = g_i(x)$ as *self-inverting*. If the transform is self-inverting and the basis functions are linearly independent, then the transform is *orthonormal* (i.e., orthogonal and normalized). One can, however, construct transforms that are self-inverting but not orthonormal, as we will discuss later.

Often, one is concerned with the properties of the individual basis and/or projection functions. In particular, signal analysis problems often require the use of functions that are spatially localized or "tuned". The property of localization implies a localization measure: for example, a function that is localized in space might have a restricted region of support, or a finite variance (second moment). Similarly, one can consider functions that are localized in the Fourier domain. A function that has a localized Fourier spectrum is tuned for a particular range of scales. Shifting this spectrum along the log Fourier axis corresponds to a dilation of the function in the spatial domain.

In two dimensions, the Fourier domain may be parameterized in terms of polar coordinates. Intuitively, this is a more natural coordinate system than the Cartesian system: the (logarithmic) radial component corresponds to scale, and the angular coordinate corresponds to orientation. A two-dimensional function whose Fourier spectrum is localized in these parameters is tuned for scale and orientation, and translation in this two-dimensional (log polar) Fourier domain corresponds to dilation and rotation of the function in the spatial domain.

A function that is simultaneously tuned for several parameters is said to be *jointly localized* in those parameters. The concept of joint localization in space and spatial frequency was introduced by Gabor in 1946 [16]. He derived a lower bound for a particular measure of joint localization, and determined the unique set of functions (products of sinusoids and Gaussians) that achieve this limit[1]. More recently, Daugman extended the Gabor basis set to two dimensions, and emphasized the importance of analyzing orientation in images [11].

Many authors have worked on the analysis of orientation (e.g., Knutsson and Granlund [22]). Freeman and Adelson [14, 15] developed transforms for analysis of oriented structure in images. They developed a technique to synthesize kernels of arbitrary orientation from linear combinations of a fixed set of oriented basis kernels, and they called these basis sets "steerable" filters. They describe the conditions under which a given set of rotated basis functions is steerable. Perona [35] has described an alternative formulation of the steerability property and recently extended this to the analysis of scale.

Multi-scale analysis has been another important topic in signal processing. The basis functions of Gabor's original transform were chosen to cover equal-width frequency bands. These functions are therefore related by modulation in the frequency domain. In order to create representations more suited for analysis of scale, many authors in signal processing and computer vision as well as biological vision have advocated transforms with constant logarithmic-width (constant Q) frequency bands [19, 31, 48, 24, 1, 34, 28]. The basis functions of such a transform are related by dilation operations. Burt and Adelson [5] developed an

[1] Lerner [25] later noted that the choice of a different localization measure could result in a different class of "optimal" function, casting some doubt on the unique importance of Gabor's class of functions.

efficient octave-width subband transform called the "Laplacian pyramid" and used it for image coding and other image processing tasks. An advantage of their technique is the computational efficiency of the recursive pyramid structure.

In an independent context, mathematicians developed a form of continuous function representation called *wavelets* [20, 9, 10]. The defining characteristic of a wavelet transform is that the basis functions are translations and dilations of a common kernel (see [42] for an introduction). Thus, the wavelet transform performs a scale decomposition similar to the Laplacian pyramid. Furthermore, it may be implemented in a recursive manner similar to the Laplacian pyramid. Finally, the term wavelet is usually assumed to refer to an orthonormal basis set.

Prior to the development of wavelets, octave-width subband transforms based on quadrature mirror filters (QMFs) were developed in the signal processing community [7, 13, 45, 44, 46, 39, 29]. In the discrete domain, pyramids constructed from QMFs may be considered as wavelet transforms, although they are often only approximately orthonormal. They have been used for various tasks requiring multi-scale signal analysis, and have been found especially effective for data compression. In two dimensions, separable implementations have been used for image compression [50, 17, 3, 43]. QMFs with oriented basis functions that are related by translations, dilations and rotations have been developed on hexagonal sampling lattices [3, 40]. Several other authors have developed non-orthogonal oriented multi-scale transforms in which the basis functions are translations and rotations of a common function [12, 19, 32, 47].

As mentioned in the introduction, one major drawback of using wavelets (or QMFs) for signal processing applications is their lack of translation invariance [29, 42]. In the next section, we address this problem by defining *shiftability*, a generalization of both the sampling theorem and the steerability property introduced by Freeman and Adelson [14, 15]. We discuss this property in the context of different image transform parameters: position, scale, and orientation. To achieve shiftability, we show that one must relax the orthogonality property of the wavelet transform. We do not, however, need to abandon the property of self-inversion.

The concept of joint localization introduced by Gabor motivates the decomposition of images in terms of spatial position, scale, and orientation. However, this does not guarantee that these parameters are easily accessible. The concept of shiftability that we introduce here emphasizes the sampling and interpolation of this parameter space, and the ability to independently access these parameters.

3 Shiftability

In this section, we define the property of transform shiftability and discuss its application to the domains of spatial position, orientation, and scale.

Suppose that the basis functions of a two-dimensional transform are translations, dilations, and rotations of a common function (kernel). Suppose also that this kernel is simultaneously localized in spatial position, scale and orientation. Then each basis function may be associated with a point in the space of parameters, (x, y, r, θ), where x and y correspond to the spatial location and r and θ correspond to scale and orientation (the log polar coordinates of the Fourier domain). Given the transform coefficients associated with these basis functions, we want to be able to compute coefficient values corresponding to *any* point in this parameter space. Although the four parameters may assume continuous values, we do not want an infinite

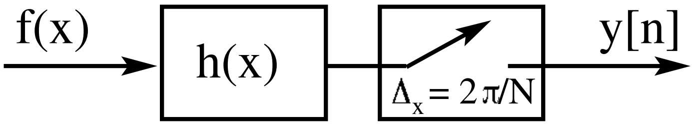


Figure 2: Block diagram illustrating the filtering and sampling operation

set of basis functions, each tuned for slightly different parameter values. Instead, we want to *interpolate* responses corresponding to any point in parameter space from the coefficients corresponding to a finite set of samples.

Moreover, we would like to continuously vary a single parameter while leaving the others fixed. For example, for a fixed subband (i.e., fixed orientation and scale), one should be able to interpolate subband responses at any spatial position from the coefficients in that subband. At a fixed location and scale, one should be able to interpolate responses at any orientation given the set of coefficients corresponding to the oriented basis functions at that location and scale. And for fixed location and orientation, one should be able to compute responses at any scale, given the coefficients of basis functions of different scales at that location and orientation.

Note that interpolation is always possible when the transform is invertible: one can simply invert the transform and then compute the inner product of the signal with the appropriate projection function. For example, one could interpolate between samples in one subband of a wavelet transform by reconstructing the original signal, and then applying the appropriate projection function at the appropriate position. But one cannot in general interpolate the highpass response using only the highpass coefficients, due to the aliasing that occurs during the subsampling operation.

Shiftability in Position

To define shiftability more concretely, we discuss it in a context that is familiar to most readers. Suppose one wishes to analyze a one-dimensional signal $f(x)$ by convolving it with a linear kernel $h(x)$. The convolution output is to be sampled: that is, the convolution output is not to be retained at every position. We do not, however, want the sampling operation to affect the representation (i.e., the points at sampled locations should not be treated differently than other locations in the original signal). Specifically, one should be able to compute (interpolate) the full convolution output from the set of retained samples. It is well-known that the lower bound on the sampling rate that meets this condition is the Nyquist rate: the frequency bandwidth of the kernel $h(x)$ [33]. Thus, in the spatial domain, shiftability of a sampled representation corresponds to the lack of aliasing.

We derive an alternative form of the sampling theorem that makes explicit a set of analytic *interpolation* functions that can be used to perform the shifting operation. The following discussion is restricted to continuous input signals, $f(x)$, that can be expanded in a Fourier series $F(k)$ (i.e., periodic or compactly supported signals); the result is easily extended to discretely sampled signals. For notational simplicity, assume that the signal period is 2π. A set of N transform coefficients, $y[n]$, are computed by convolving with a kernel $h(x)$ and uniformly sampling the output[2] with sample spacing $\Delta_x = \frac{2\pi}{N}$:

$$y[n] = \int_0^{2\pi} dx\, h(n\Delta_x - x) f(x), \qquad n \in \{0, 1, \ldots N-1\}. \tag{3}$$

This is illustrated in the block diagram in figure 2. In this linear system, the projection functions correspond to a set of shifted copies of the kernel: $\{h(n\Delta_x - x) \mid n = 0, 1, \ldots, N-1\}$.

The transformation described in equation (3) is shiftable if there exists a set of interpolation functions, $b_n(x_\circ)$, that can be used to compute the inner product of the signal with an arbitrarily shifted version of the kernel $h(x)$ as a weighted combination of the $y[n]$:

$$\int_0^{2\pi} dx\, h(x_\circ - x) f(x) = \sum_{n=0}^{N-1} b_n(x_\circ) y[n] \tag{4}$$

where $x_\circ$ is the arbitrary shift distance. The following proposition (adapted from [15]) describes a condition under which the transformation is shiftable, and gives an analytic expression for the interpolation functions.

Proposition 1 *The transformation defined by equation (3) is shiftable if and only if there exist a set of interpolation functions, $b_n(x_\circ)$, that satisfy the matrix equation*

$$\begin{pmatrix} e^{jx_\circ k_0} \\ e^{jx_\circ k_1} \\ \vdots \\ e^{jx_\circ k_{M-1}} \end{pmatrix} = \begin{pmatrix} 1 & e^{j\Delta_x k_0} & e^{j2\Delta_x k_0} & & e^{j(N-1)\Delta_x k_0} \\ 1 & e^{j\Delta_x k_1} & e^{j2\Delta_x k_1} & \cdots & e^{j(N-1)\Delta_x k_1} \\ & \vdots & & & \\ 1 & e^{j\Delta_x k_{M-1}} & e^{j2\Delta_x k_{M-1}} & & e^{j(N-1)\Delta_x k_{M-1}} \end{pmatrix} \begin{pmatrix} b_0(x_\circ) \\ b_1(x_\circ) \\ \vdots \\ b_{N-1}(x_\circ) \end{pmatrix} \quad \text{for all } x_\circ, \tag{5}$$

where $\{k_0, k_1, \ldots, k_{M-1}\}$ is the set of frequencies for which $H(k)$, the Fourier series of $h(x)$, is non-zero.

Proof: We start with the definition of shiftability given in equation (4):

$$\begin{aligned} \int_0^{2\pi} dx\, h(x_\circ - x) f(x) &= \sum_{n=0}^{N-1} b_n(x_\circ) y[n] \\ &= \sum_{n=0}^{N-1} b_n(x_\circ) \int_0^{2\pi} dx\, f(x) h(n\Delta_x - x) \\ &= \int_0^{2\pi} dx\, f(x) \sum_{n=0}^{N-1} b_n(x_\circ) h(n\Delta_x - x). \end{aligned}$$

In order for this expression to hold for any signal $f(x)$, it must be true that

$$h(x_\circ - x) = \sum_{n=0}^{N-1} b_n(x_\circ) h(n\Delta_x - x). \tag{6}$$

That is, we can write the arbitrarily shifted kernel $h(x - x_\circ)$ as a linear combination of the basis functions $h(x - n\Delta_x)$ that are positioned at intervals of the sample spacing, Δ_x. This

[2]The following derivation does not rely on the uniformity of the sampling, but the translation invariance of the coefficient power (discussed in the next section) only holds in the case of uniform sampling.

means that the sampled basis set spans the entire subspace of all translations of the kernel. Equivalently, the linear subspace spanned by the sampled basis set is invariant to translation.

Since we are assuming $f(x)$ is periodic, we may assume without loss of generality that $h(x)$ is also periodic. Let $H(k)$ be the (complex) Fourier series of $h(x)$. Then flipping the sign of x and computing the Fourier series of both sides of equation (6) produces:

$$H(k)e^{jkx_\circ} = H(k)\sum_{n=0}^{N-1} b_n(x_\circ)e^{jkn\Delta_x}, \qquad \text{for all } x_\circ, k. \tag{7}$$

This is a set of linear constraints on the interpolation functions, $b_n(x_\circ)$, describing the phase-shifting of the individual Fourier components of $h(x)$ necessary to translate the entire function. If a solution to this equation exists, then the representation is shiftable.

Clearly, equation (7) imposes a constraint on the interpolation functions only for those frequencies k where $H(k)$ is non-zero. If $H(k)$ is non-zero for a finite set of frequencies (or if $f(x)$ and $h(x)$ are discrete signals), this constraint may be written using matrix notation. Let $\{k_0, k_1, \ldots, k_{M-1}\}$ be the set of frequencies for which $H(k)$ is non-zero. Then equation (7) may be written as the matrix equation (5) given in the proposition. When the matrix in the equation is invertible, we can solve for the interpolation functions $b_n(x_\circ)$, and the representation will therefore be shiftable. ∎

The relationship of the constraint in equation (5) to the Nyquist criterion should be clear. Invertibility of the matrix places a lower bound on the number of samples, N, required for shiftability. In particular, the number of samples must be greater than or equal to the number of non-zero Fourier components of the function $h(x)$, as discussed in [15]. For example, consider one stage of a discrete dyadic wavelet transform. If the original signal is of length (period) N samples, the highpass subband computed from this signal will contain $N/2$ samples. The kernels are designed to satisfy the constraint:

$$\left|H(k)\right|^2 + \left|H(k+\tfrac{N}{2})\right|^2 = 2, \qquad \text{for each } k.$$

Since the kernels are typically not ideal (sinc) functions, they must contain non-zero frequency components corresponding to more than $N/2$ frequencies, and therefore the matrix in equation (5) will not be invertible.

It is important to realize that except for the specification of which Fourier coefficients are non-zero, the interpolation functions are *independent of the kernel* $h(x)$. This means that if two kernels have the same set of non-zero Fourier components, they share a common set of interpolation functions.

Note that if the matrix in equation (5) is not invertible, we may still find a least-squares solution using the matrix pseudo-inverse. On the other hand, if the representation is overcomplete, then the interpolation functions will not be unique. In this case, we can again solve for the $b_n(x_\circ)$ using the pseudo-inverse (thus producing a minimum-norm set of interpolators). Alternatively, one may prefer to use a set of interpolators with another property such as minimal region of support.

Shift-invariant Response Power

In an aliased transform, the response power depends on the signal position: translation of the input signal generally results in a re-distribution of the power content amongst the various

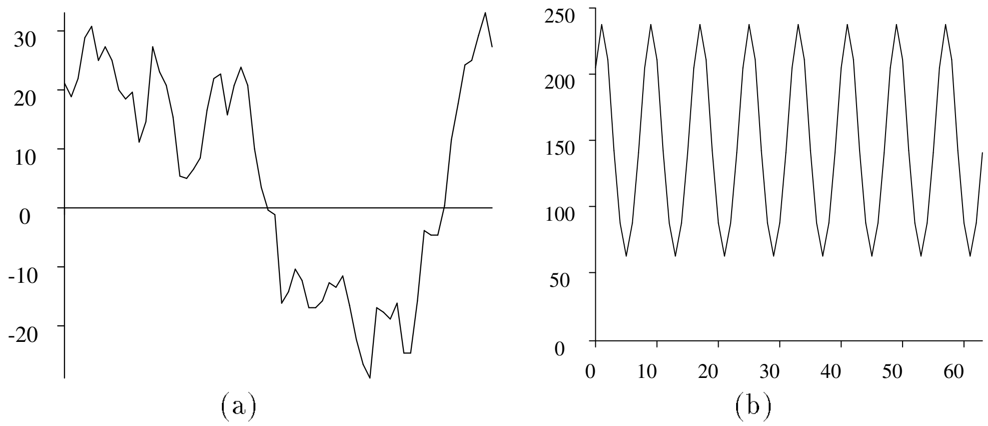


Figure 3: Illustration of the dependency of subband power on signal position. (a) A discrete fractal signal, of length 64. We constructed a three-level wavelet pyramid decomposition of this signal, using the four-tap Daubechies wavelet. This was done for 64 (circularly) translated copies of the input signal. (b) Plot of the power in the third wavelet subband as a function of translation distance of the input signal. Note that the power varies substantially as the input is shifted, although translation produces no change in the spectral power of the signal.

frequency subbands, even though there has been no change in the frequency spectrum of the input. This is illustrated in figure 3. We show a fractal signal, and the power (sum of square coefficients) in the third level subband of a discrete dyadic wavelet expansion as a function of signal position. As the signal is (circularly) shifted, the power in the subband oscillates over a wide range of values.

The shiftability constraint is equivalent to the constraint that the power of the transform coefficients in the subband is preserved when the input signal is shifted in position. We show this in the following proposition:

Proposition 2 *Given a set of transform coefficients computed using equation (3), the transform is shiftable if and only if the power of the coefficients, $\sum_{n=0}^{N-1} |y[n]|^2$, is invariant to translations of the input signal $f(x)$, for any choice of input signal.*

Proof: We show this result by considering the discrete Fourier transform (DFT), $Y(k)$, of the output of the diagram in figure 2. Using standard facts about sampling, we may write

$$Y(k) = \sum_l H(k + lN) \cdot F(k + lN), \tag{8}$$

where $F(k)$ is the Fourier series of the input signal, $H(k)$ is the Fourier series of the convolution kernel, and the subscript l ranges over the set of integers. The summation corresponds to the spectrum replication that occurs when the signal is sampled.

Assume that the transform is shiftable. Using Parseval's theorem, we write the power of the output signal in terms of its DFT:

$$\begin{aligned}
\sum_{n=0}^{N-1} |y[n]|^2 &= \sum_{k=0}^{N-1} |Y(k)|^2 \\
&= \sum_{k=0}^{N-1} \left| \sum_l H(k + lN) \cdot F(k + lN) \right|^2
\end{aligned}$$

$$= \sum_{k=0}^{N-1} \sum_{l} |H(k+lN) \cdot F(k+lN)|^2$$
$$+ \sum_{k=0}^{N-1} \sum_{l \neq m} H(k+lN) \cdot F(k+lN) \cdot H^*(k+mN) \cdot F^*(k+mN). \quad (9)$$

where the $()^*$ operator is complex conjugation, and the subscripts l and m range over all integers. Since we are assuming shiftability (i.e., no aliasing), the second term is zero. The remaining expression is invariant to translations of the input signal:

$$\sum_{n=0}^{N-1} |y[n]|^2 = \sum_{k=0}^{N-1} \sum_{l} |H(k+lN)|^2 |F(k+lN)|^2 \quad (10)$$

That is, replacing $F(k)$ by $e^{jkx_\circ}F(k)$ for any $x_\circ$ will not alter the right hand side of expression in equation (10): the reponse power is translation invariant.

The converse is also true. Assume the system is not shiftable: then the sampling causes aliasing. Then a sinusoidal input signal at a frequency that is aliased will appear in the second term of equation (9) (i.e., there will be some k, l, and m for which the product in the second term is non-zero). As the input signal is translated, this non-zero second term will be modulated by a phase factor. ∎

In situations where it is impossible or impractical to achieve shiftability with respect to a parameter, there is a more easily satisfied constraint that may be used in its place. We call this property "flat basis power" with respect to the parameter. In the case of the spatial position parameter, this means that the sum of the power of the basis functions for one subband is constant as a function of position. This property is a consequence of the shiftability property, but the converse is not true. We will use this constraint in the design of a two-dimensional transform in section 5. Here, we define the property and show that it a direct consequence of shiftability.

Corollary 3 *If the transform specified by equation (3) is shiftable, the sum of squares of the basis functions, $\sum_n h_n^2(x)$, is constant for all x.*

Proof: We show this result by considering the impulse input signal ($F(k) = e^{j\omega x_\circ}$). Then $y[n]$ will be the value of the sample of the nth basis function at the location of the impulse. In this case, the left hand side of equation (10) is equal to the sum of squares of a subset of the kernel samples spaced by the sampling distance:

$$\sum_{n=0}^{N-1} |y[n]|^2 = \sum_{n} |h(x_\circ - n\Delta_x)|^2$$

The right hand side of equation (10) is the sum over all frequencies of the kernel power spectrum: a constant. Thus, the sum of squares of the basis functions at any location must equal the same constant:

$$\sum_{n} |h(x - n\Delta_x)|^2 = c, \qquad \text{for all } x. \quad (11)$$

∎

We emphasize that flat basis power of the transform is a necessary but not a sufficient condition for shiftability. For example, consider convolution with the interval integration kernel,

$$h(x) = \begin{cases} 1, & x \in [0,1], \\ 0, & \text{otherwise,} \end{cases}$$

followed by sampling at unit intervals. This system will clearly have flat basis function power, since the power of each basis function is one over its interval, and the functions do not overlap. But this set is not shiftable, since the kernel spectrum is infinite in extent and therefore, the sampling of the basis set introduces aliasing.

Shiftability in Orientation

We can apply the concept of shiftability to the continuous shifting of orientation. Given a set of basis functions centered at the same spatial location and scale, and tuned to a finite number of orientations, we wish to interpolate measurements at all orientations. This problem has been studied in detail by Freeman and Adelson [15], who described a general framework for continuously rotating filters. They discuss the number of rotated basis functions required to "steer" a given filter (i.e. the number of samples in the orientation domain), and derive the interpolation formula given in equation (5).

To illustrate this idea, consider the set of directional derivatives of a two-dimensional radially symmetric function, $c(r)$, where $r = \sqrt{x^2 + y^2}$ is the radial coordinate. The directional derivative of this function at orientation $\theta_\circ$ is written as:

$$D_{\theta_\circ}[c(r)] = \cos(\theta - \theta_\circ)c'(r),$$

where θ is the angular coordinate and $c'(r) = \frac{d}{dr}[c(r)]$, the derivative of $c(r)$ with respect to its argument. Note that this function is polar separable, a product of angular and radial component functions. Using standard trigonometric relations, we can rewrite this expression in the form of the shiftability constraint in equation (6):

$$\begin{aligned} \cos(\theta - \theta_\circ)c'(r) &= \cos(\theta_\circ)\cos(\theta)c'(r) + \sin(\theta_\circ)\sin(\theta)c'(r) \\ &= \cos(\theta_\circ)\left[\cos(\theta)c'(r)\right] + \sin(\theta_\circ)\left[\cos(\theta - \tfrac{\pi}{2})c'(r)\right]. \end{aligned}$$

Thus, the rotation of the directional derivative to an arbitrary angle, $\theta_\circ$, may be computed as a linear combination of the directional derivative basis functions at angles 0 and $\frac{\pi}{2}$ (in brackets). The basis set composed of these two derivatives is shiftable in orientation.

We can also consider steerability in the Fourier domain: That is, the steerability of the *Fourier transforms* of the basis functions. The steerability condition is easiest to state for polar-separable functions. Assume that the frequency domain function $V(\omega_x, \omega_y)$ may be written separably in polar coordinates as

$$\tilde{V}(r, \theta) = H(\theta)U(r),$$

where r and θ are the polar coordinates of the frequency domain. $H(\theta)$ is clearly periodic (with period 2π), and shifting this function with respect to its argument corresponds to rotating the two-dimensional basis function $V(\omega_x, \omega_y)$. Thus, shiftability of $H(\theta)$ corresponds to "steerability" of $V(\omega_x, \omega_y)$. The shiftability constraint is now written in terms of the Fourier components of the function $H(\theta)$. Assuming that the transform basis set is composed of rotations of this function at a sufficient number of angles, the interpolation (steering) functions may be derived from equation (5).

In general, we may also consider orientation shiftability of functions that are not polar-separable. We write the function in terms of a family of angular functions, $H_r(\theta)$:

$$\tilde{V}(r, \theta) = H_r(\theta)U(r).$$

Recall that the shiftability constraint in equation (5) depends only on which frequency components of the function are non-zero; it does not depend on the magnitude of those components. Therefore, we can express the shiftability constraint in terms of the union of the non-zero frequency components of $H_r(\theta)$ for all r.

Analogous to the case of position shiftability, orientation shiftability ensures that the sum of the squares of the transform coefficients at a given location and scale will be invariant under changes in the input image orientation. Consider, for example, a sinusoidal input of spatial frequency $\vec{\omega}_\circ$, with unit amplitude. If we rotate the input about a point (i.e., modify the angular coordinate of $\vec{\omega}_\circ$), the response power (sum of squares of the coefficients) at that point will not change.

Shiftability in Scale

Application of the shiftability concept to scale is more difficult both because of the symmetry constraints, and because we want to express the parameter on a logarithmic axis. For simplicity, we consider functions that are purely symmetric or anti-symmetric about the origin in the frequency domain. We handle this symmetry constraint by constructing a *reflection-shiftable* function.

Proposition 4 *Consider a shiftable representation based on function $H(r)$, with interpolation functions $b_n(r_\circ)$. Then we can compute the function $U_\pm(r, r_\circ) = H(r - r_\circ) \pm H(-r - r_\circ)$ from the set of functions $U_\pm(r, n\Delta_r) = H(r - n\Delta_r) \pm H(-r - n\Delta_r), n \in \{0, 1, \ldots, N-1\}$, using the same interpolation functions. We say that $U_\pm(r, r_\circ)$ is* reflection-shiftable.

Proof: The proof is simple. We assume $H(r)$ is shiftable:

$$H(r - r_\circ) = \sum_n b_n(r_\circ)H(r - n\Delta_r).$$

Then

$$\begin{aligned}
U_\pm(r, r_\circ) &= H(r - r_\circ) \pm H(-r - r_\circ) \\
&= \sum_n b_n(r_\circ)H(r - n\Delta_r) \pm \sum_n b_n(r_\circ)H(-r - n\Delta_r) \\
&= \sum_n b_n(r_\circ)\left[H(r - n\Delta_r) \pm H(-r - n\Delta_r)\right] \\
&= \sum_n b_n(r_\circ)U_\pm(r, n\Delta_r).
\end{aligned}$$

This is the desired result. [3] ∎

Typically, signals of interest will be band-limited, and thus we can limit our domain to a particular range: $r \in [-\omega_{\max}, \omega_{\max}]$. Since we have developed our discussion of shiftability in the context of periodic signals, we can make the domain periodic by identifying the endpoints. We will use this approach in designing a scalable transform in section 4.

The only remaining problem is that the reflection-shiftable kernels described above correspond to equal-width frequency subbands. Consider a set of basis functions corresponding to dilations of a common kernel. The fourier transforms of these basis functions will be shifted copies of the kernel transform *in log frequency.* Thus, we would like to have frequency subbands that are of equal size on a *logarithmic* axis. To accomplish this, we note that the shiftability property is not affected by "warping" the domain. For example, if we assume that

$$U(r, r_\circ) = \sum_n b_n(r_\circ) U(r, n\Delta_r),$$

then clearly

$$U\left(\rho(r), r_\circ\right) = \sum_n b_n(r_\circ) U\left(\rho(r), n\Delta_r\right),$$

where $\rho(r)$ is any warping function. Note that this warping operation will not affect the flat power property (i.e., corollary 3) of the basis functions.

Since we are interested in functions that are tuned for scale, the $\rho(r) = \log(r)$ is the correct warping function. In this case, the singularity at the origin presents a problem and so it is necessary to modify the function near the origin. This does not affect the shiftability of the basis functions; it does, however, mean that they are not precisely related by dilation operations. An example (modified) log function is illustrated in figure 4. In section 4, we develop a one-dimensional "scalable" (i.e., shiftable in scale) transform using a logarithmic warping function.

Joint Shiftability

In the previous sections, we considered shiftability of several transform parameters independently. It is natural to ask whether a transform can be simultaneously shiftable in several parameters. By this, we mean the representation is shiftable in each parameter while all other parameters are held fixed. For different applications, it may be desirable to design transforms that are shiftable in different subsets of parameters. We discuss some of the basic issues in joint shiftability; a full analysis of the topic is beyond the scope of this paper.

We first note that for parameters that are independent, joint shiftability can be achieved. For example, one can create a two-dimensional transform that is shiftable in both x and y position. These two parameters may be treated independently, and the shifting may be performed separably. As an example of this, in section 5 we discuss the implementation of a transform that is jointly shiftable in two-dimensional position and orientation.

If, however, we consider parameters that are Fourier complements of each other, there are difficulties. For example, consider a one-dimensional transform with basis functions parameterized for position and scale. Position shiftability (with spatial subsampling) requires the basis

[3] Note that the flat basis power condition described in corollary 3 will only be met if we include both the symmetric and the anti-symmetric basis functions when computing the power. That is, the sum $\sum_n \left[U_+(r, n\Delta_r)^2 + U_-(r, n\Delta_r)^2\right] = 2\sum_n \left[H(r - n\Delta_r)^2 + H(-r - n\Delta_r)^2\right]$ will be constant if $H(r)$ is shiftable.

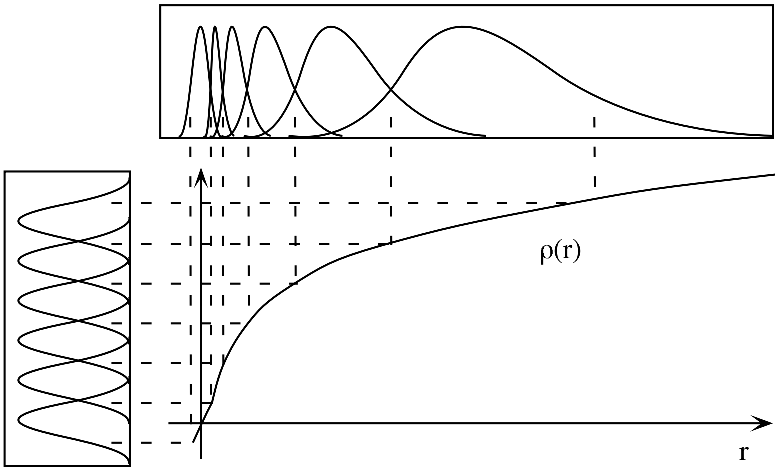


Figure 4: Illustration of the use of a warping function. The function is a modified logarithm: for argument values near the origin it is linear. Along the top are the set of warped subbands. The kernels corresponding to these subbands are only approximately related by dilations: the modification near the origin prevents them from being exact. Along the left is a shiftable set of functions. The same interpolation functions for the shiftable set may be used to shift the warped functions along the warped frequency axis.

functions to have limited bandwidth (regions of support) in the Fourier domain. This would imply that the basis functions had infinite spatial regions of support in the spatial domain. On the other hand, scale shiftability (with scale subsampling) requires them to have compact regions of support in the spatial domain. Thus it is impossible to design a transform that is subsampled in two Fourier complementary domains, and shiftable in both of those domains.

It should now be clear that the critical sampling of orthogonal wavelet transforms prevents them from being shiftable in both position and scale. If we use an ideal (sinc) kernel in a dyadic wavelet transform, the subbands will be spatially shiftable. But they will be infinite in spatial extent and highly non-shiftable in scale.

There are several possible solutions to this dilemma. The simplest is to maintain full resolution in one of the parameters. In this case, there are no compactness restrictions on the complementary parameter. For subband transforms, this inefficiency may be unacceptable for many applications. Note, however, that the Fourier basis itself may be considered to be jointly shiftable in this sense. For each frequency ω there are two spatially translated basis functions ($\sin(\omega x)$ and $\cos(\omega x)$). Each of these pairs is shiftable (one can compute a sinusoid of this frequency at any phase from these two basis functions). The transform is maintained at full density in the frequency domain (by definition), and so is trivially shiftable there.

This example also suggests a connection between the shiftability concept and eigensystem analysis. The Fourier basis functions are the eigenfunctions of the translation operator, when the basis functions are considered as complex pairs. This means that the two-dimensional subspace spanned by the two basis functions of a given frequency is invariant under translations. Analogously, a set of basis functions shiftable in a parameter forms a multi-dimensional eigenfunction for translations with respect to that parameter. The space spanned by these basis functions is invariant to modifications of the parameter.

Returning to the issue of joint shiftability, one can also consider non-separable shiftability, in which the interpolation of a point in the parameter space is based on a set of sample points not restricted to variations of a single parameter. In this case, the interpolation functions are multi-dimensional. This is true, for example, in a wavelet transform: One can interpolate any point in the scale-position parameter space from a set of surrounding scale and position samples. This was done in a recent paper by Gopinath and Burrus [18]. In the present paper, however, we have emphasized the advantages of being able to shift independently, so that one parameter may be shifted while holding all others fixed. This property is important in many signal analysis applications, as we discuss in later sections.

Another alternative is to design an approximately shiftable representation. This would require introduction of a measure of joint nonshiftability (joint aliasing); basis functions could then be designed to minimize this measure, given a fixed sampling structure. Note that such a measure would necessarily be a measure of joint localization, as in Gabor's work. But since it would be based on region of support, it would lead to different sets of "optimal" basis functions. Note also that this measure would be dependent on the sampling density chosen: the joint shiftability constraint becomes less restrictive as one increases the number of transform samples.

Finally, given that joint shiftability is impossible to achieve in a pair Fourier complementary domains, we can also consider relaxing the shiftability constraint in one of the domains. In particular, we can replace the shiftability constraint in one domain by the less restrictive constraint that the basis function power is flat (see corollary 3). In the next section, we show that this will result in a self-invertible transform.

Self-Invertibility

Orthonormal transforms are self-inverting, but self-inverting transforms need not be orthonormal. As we stated in section 2, orthonormality is the combination of self-inversion *and* linear independence. One can construct transforms which are self-inverting, yet overcomplete. Daubechies has termed these "tight frames" [8].

A trivial example serves to demonstrate that such transforms exist. A non-orthogonal, self-inverting transform may be formed from any orthogonal transform by duplicating every projection function and dividing the kernel weights by a factor of $\sqrt{2}$. The result is an overcomplete representation in which each pair of identical projection functions encodes the same information. But the transform remains self-inverting: the basis and projection functions are identical. A less trivial example of an overcomplete self-inverting transform based on an overcomplete Gabor set is described by Simoncelli [39]. The "cortex" transform developed by Watson [47] is another example of an overcomplete self-inverting transform. The overcomplete steerable pyramid described in section 5 is also a self-inverting transform.

Intuitively, the importance of self-invertibility is that it ascribes a natural meaning to the transform coefficients. Each coefficient is computed using a projection function with a particular position and shape in both the spatial and frequency domains. The basis functions used to invert the transform are the same. This can be important for many applications, such as data compression and image enhancement. Without this property, errors introduced by nonlinear processing of the coefficients will spread to locations and frequencies other than those that were used to compute the coefficients.

In previous sections, we have concentrated on one transform parameter at a time. Now we

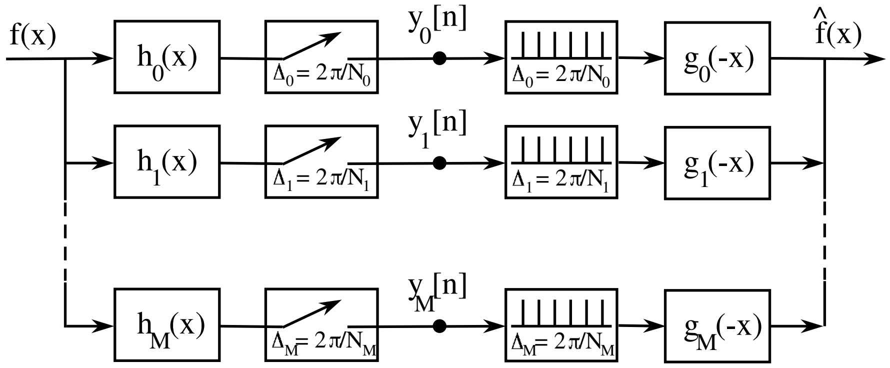


Figure 5: An analysis/synthesis filter bank.

must consider the transform as a whole. We restrict our attention to the class of transforms that may be described by analysis/synthesis filter banks, as illustrated in figure 5. In fact, this structure is general enough to represent any linear transform, but we will have in mind a subband transform such as the wavelet transform.

As before, we will assume that the input signal, $f(x)$, is periodic with period 2π. In addition, we will need to assume that the input signal is band-limited to a finite range of frequencies, $k \in [-k_{\max}, k_{\max}]$. In the system depicted in figure 5, the input signal is convolved with a set of M kernels, $h_m(x)$. The index, m, corresponds to the frequency subband of the kernel. The output of each convolution is sampled at N_m points at intervals of $\frac{2\pi}{N_m}$. This produces a set of transform coefficient sequences, $y_m[n]$.

In the synthesis section, the coefficients are reconstituted as a continuous signal by attaching them to Dirac delta functions at the proper sample locations. That is, from the discrete signal $y_m[n]$, we form a continuous signal $\gamma_m(x) = \sum_{n=0}^{N_m-1} y_m[n] \cdot \delta(x - n\Delta_x)$. These signals are then convolved with the space-reversed kernels, $g_m(-x)$. The negation of the argument allows us to be consistent about the definitions of basis and projection functions. Given this structure, we can show the following:

Proposition 5 *A subband transform that is spatially shiftable in each subband and exhibits flat basis function power in the complementary Fourier domain over all frequencies of interest is self-inverting.*

Proof: The proof is by construction in one dimension; the extension to higher dimensions is straightforward. We assume that the conditions of the proposition hold, and that the projection filters are equal to the basis filters: $g_m(x) = h_m(x)$. Then we must show that the system in figure 5 is an identity system.

Using equation (8), the output of the entire system may be written in the frequency domain as:

$$\begin{aligned}
\hat{F}(k) &= \sum_{m=0}^{M-1} \left[\sum_{l=0}^{N_m-1} H_m\left(k + lN_m\right) \cdot F\left(k + lN_m\right) \right] H_m(-k) \\
&= \sum_{m=0}^{M-1} H_m(k) H_m(-k) \cdot F(k)
\end{aligned}$$

$$+\sum_{m=0}^{M-1}\sum_{l=1}^{N_m-1} H_m\left(k+lN_m\right)H_i(-k)F\left(k+lN_m\right). \tag{12}$$

The second sum in this equation is the aliasing term. Since we are assuming that the system is shiftable (aliasing-free) in the spatial domain, the second sum evaluates to zero. The remaining (shift-invariant) system response is then

$$\hat{F}(k)=\left[\sum_{m=0}^{M-1}\left|H_m(k)\right|^2\right]\cdot F(k).$$

Now, if the set of $H_m(k)$ are designed to have flat power as in equation (11), then

$$\sum_{m=0}^{M-1}\left|H_m(k)\right|^2=c, \qquad \text{for all } k.$$

If we impose the additional requirement that the value of the constant c is unity (this is just an adjustment of the normalization of the basis functions), then the overall system response will be unity, and the representation is self-inverting. ∎

We will use this result in section 5 to construct a self-inverting pyramid transform that is shiftable in space and orientation.

4 A One-dimensional Example

In this section, we describe a one-dimensional transform that is approximately shiftable in scale and position and we use it in a simple example application. We designed a shiftable function, $H(r)$, by fitting an inverted parabola with one period of a five-term Fourier series. We then constructed a set of six (frequency domain) basis functions, $U(r,\frac{n\pi}{3})$, using the techniques described in the previous section. In particular, we warped the frequency axis using a modified log function:

$$\rho(r)=\mathrm{sgn(r)}\cdot\log(\alpha|\mathrm{r}|+1),$$

where α is a normalization factor $\alpha=(e^{\pi}-1)/\pi$. This modification to the log function allows us to handle the singularity at the origin. It does not effect the shiftability of the basis functions, although the basis functions will no longer be exact dilations of each other.

To form a basis set, we combined the warped functions in a reflection-symmetric manner:

$$U(r,\tfrac{n\pi}{3})=H(\rho(r)-\tfrac{n\pi}{3})+H(-\rho(r)-\tfrac{n\pi}{3}), \qquad n\in\{-2,-1,0,1,2\}.$$

Two of the basis functions, whose Fourier transforms are tuned for different frequency octaves, are illustrated in figure 6. We also show a basis function that has been interpolated (shifted) to an intermediate scale. The kernels used for convolution in the spatial domain are computed as the Fourier transforms of the $U(r,\frac{n\pi}{3})$.

Scale-Space Decomposition

To illustrate the use of this set of scalable filters discussed in the previous section, we show an efficient calculation of a "scale-space" image [48]. Ordinarily, this calculation would require

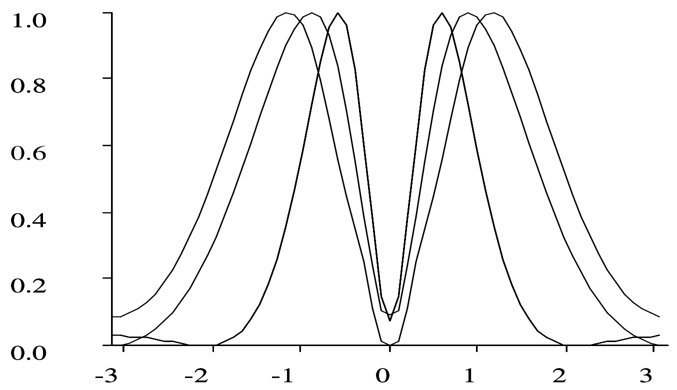


Figure 6: Real part of the frequency response of two of the scale basis functions (inner and outer curves), along with the frequency response of an interpolated basis function.

convolution of the signal with a large battery of filters, each tuned for a slightly different scale. The property of scale shiftability allows us to efficiently compute the signal at all scales, using only a small set of filters.

We first convolve a one-dimensional signal with each of our scale filters, producing a set of six one-dimensional signals $y_n(x), n \in \{0, 1, \ldots, 5\}$. As mentioned in the previous section, we must retain the full sampling density of the convolution operations in order for the subbands to be spatially shiftable. We computed a set of interpolation functions using equation (5), and interpolated a set of one-dimensional signals at intermediate scales by taking a weighted combination of the $y_n(x)$.

The results for a fractal signal are shown in figure 7. We show the original signal (a) and four of the six "basis" signals computed by convolving with each of the basis filters (b) – (e). From the set of basis signals, we can interpolate a scale signal at *any* intermediate scale. To illustrate this, we interpolated scale signals at 128 different scales ranging from that of the lowest-frequency basis scale to the highest-frequency basis scale. These were pasted together as the scan lines of a two-dimensional "scale-space" image, in which the horizontal axis corresponds to spatial position, and the vertical axis corresponds to scale. We then applied a simple zero-crossing detector to this image. Figure 7(f) shows the zero-crossings of the scale-space function. Note that the horizontal density of zero-crossings decreases as the scale becomes coarser.

5 Two-dimensional Example: A Steerable Pyramid

We have designed and implemented a two-dimensional transform that is jointly shiftable in orientation and position (see also [15]). The basis functions are translations, dilations, and rotations of a single kernel, and the transform is constructed as a recursive pyramid. Figure 8 contains an illustration of the frequency domain decomposition performed by the transform. Shiftability in scale was not needed for our intended applications, and is not a property of this transform. Our transform does, however, have flat basis power in the scale domain, thus it is self-inverting (see corollary 3). Because it is shiftable in orientation, we call it a "steerable pyramid".

Figure 9(a) illustrates the orientation-shiftable basis filters at one scale of the pyramid. The

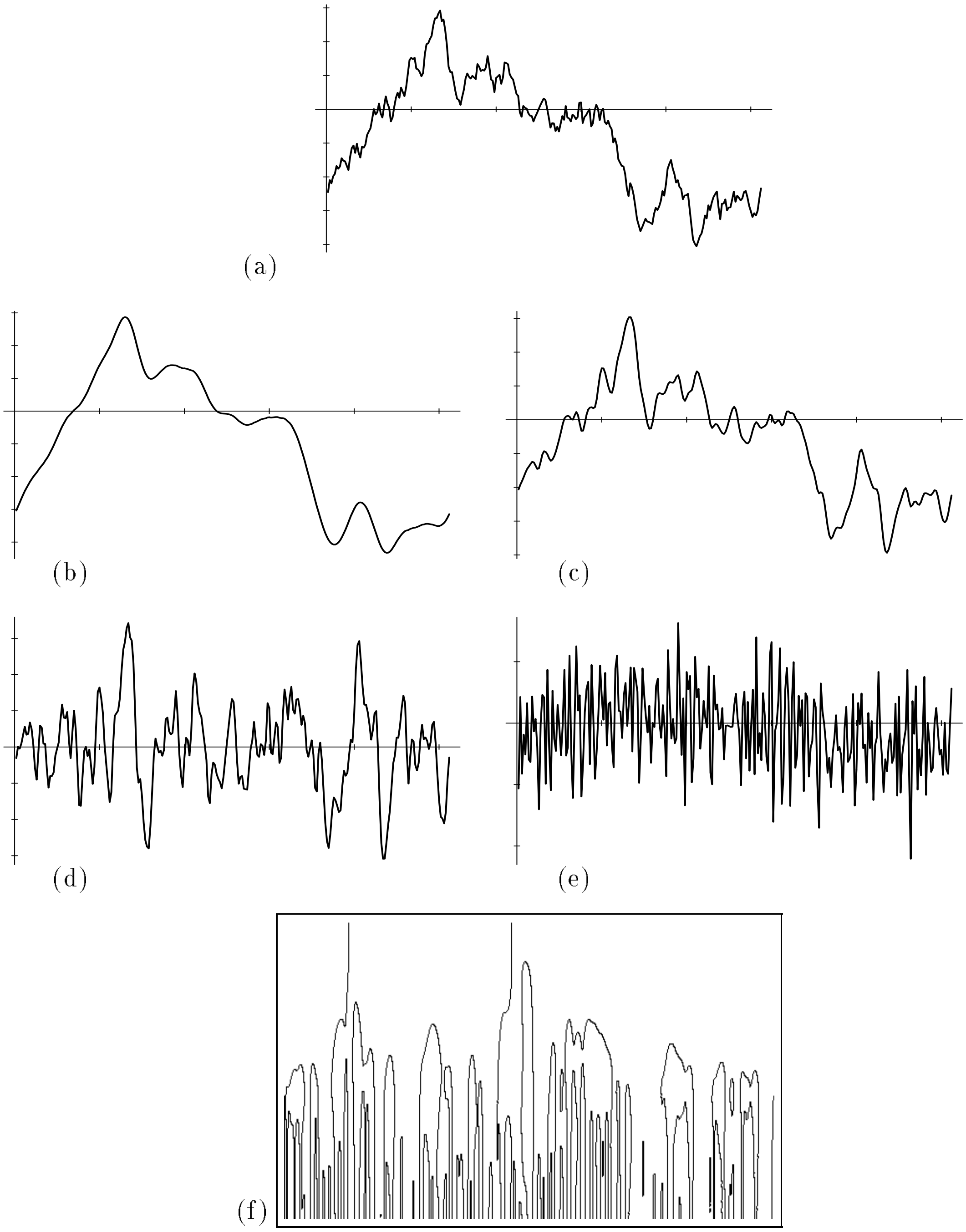


Figure 7: An example scale-space decomposition. (a) The original signal, a Brownian fractal. (b)–(e) Four of the six "scale signals" resulting from applying each of the scale basis filters to the original fractal signal. (f) The zero-crossings of the interpolated "scale-space" image (see text).

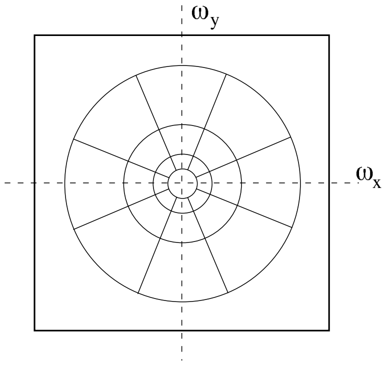


Figure 8: Illustration of the spectral decomposition performed by the steerable pyramid. The basis functions are related by dilations and rotations (except for the inner lowpass subband, and the outer residual subband).

design of the filters is described in the next section. The rest of the figure shows the decomposition of a simple test image. Unlike the orthogonal pyramids based on QMF filter banks, the steerable pyramid is significantly overcomplete: there are $\frac{16}{3}$ times as many coefficients in the representation as in the original image. The overcompleteness limits the computational efficiency but increases its usefulness for many image processing and analysis tasks, as we will show in section 6.

Pyramid (Radial Frequency) Implementation

We used a polar-separable frequency domain design strategy for the filters of the transform. In this section, we describe the radial frequency (scale) portion of the design.

Pyramid algorithms are based on recursive application of filtering and subsampling operations. Typically, the input signal is partitioned into low- and highpass portions, the lowpass portion is subsampled, and the subdivision is repeated recursively. The highpass portion may be subsampled as in wavelet pyramids, or left at full density as in the Laplacian pyramid [5]. A single stage of the transform may be written in the form of a standard analysis/synthesis filter bank [41] (although the Laplacian pyramid is typically not implemented using this architecture). The filters are chosen such that this single-stage system behaves like an identity operation. The pyramid structure is achieved by applying this single-stage transform recursively to the lowpass subband of the previous single-stage transform.

In the present case, we wish to subdivide the signal into lowpass and *bandpass* portions. To achieve this decomposition, we have implemented a novel pyramid architecture, in which the response of the single-stage system (and therefore the entire pyramid) is lowpass. The block diagram defining the pyramid recursion is shown in figure 10. The input signal is convolved with a bandpass kernel, $B(\omega)$, and a lowpass kernel $L_1(\omega)$. To ensure that there is no aliasing in the bandpass portion, it is not subsampled. The lowpass portion is subsampled by a factor of two, and then convolved with another lowpass kernel, $L_0(\omega)$.

Using standard signal processing results, and assuming that the subsampling of the lowpass branch introduces negligible aliasing, we write the response of the system of figure 10 as:

$$S(\omega) = |B(\omega)|^2 + |L_1(\omega)|^2 \, |L_0(2\omega)|^2 \tag{13}$$

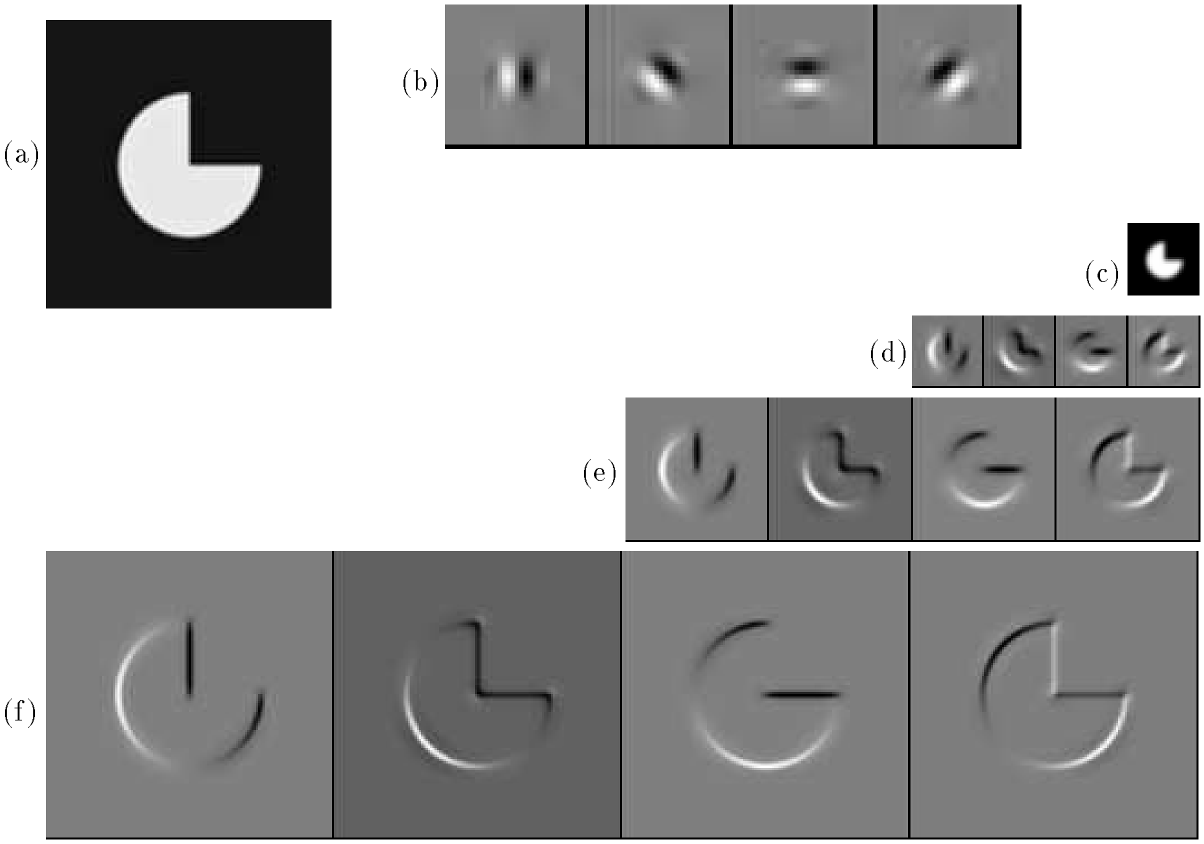


Figure 9: Illustration of the structure of the steerable pyramid. (a) A test image. (b) Filter kernels used to construct the pyramid. These four filters, which are rotated versions of a common filter, span the space of all rotations of that common filter. (c) Lowpass coefficient image in the pyramid representation of the test image. (d) – (f) Bandpass coefficient images in pyramid representation.

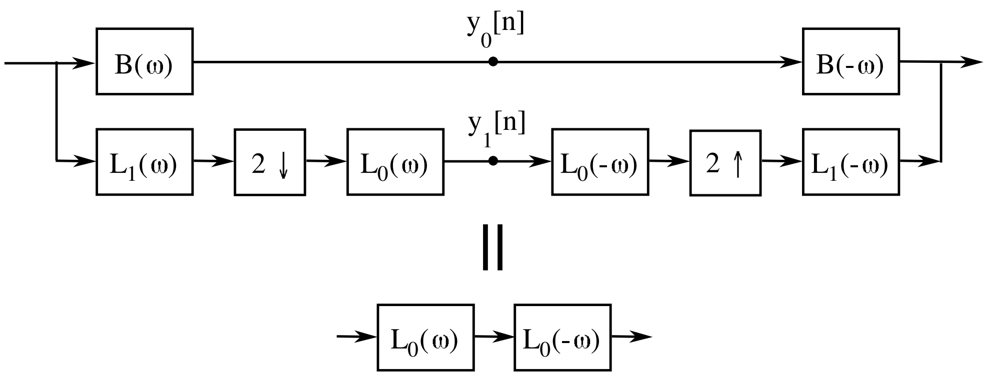


Figure 10: Illustration of the design constraint for the radial component of a steerable, self-inverting pyramid. The impulse response of the entire system (shown at the top) equals the power spectrum of the filter $L_0(\omega)$. This allows us to cascade the system recursively, as illustrated in figure 11, to create a multi-scale transform.

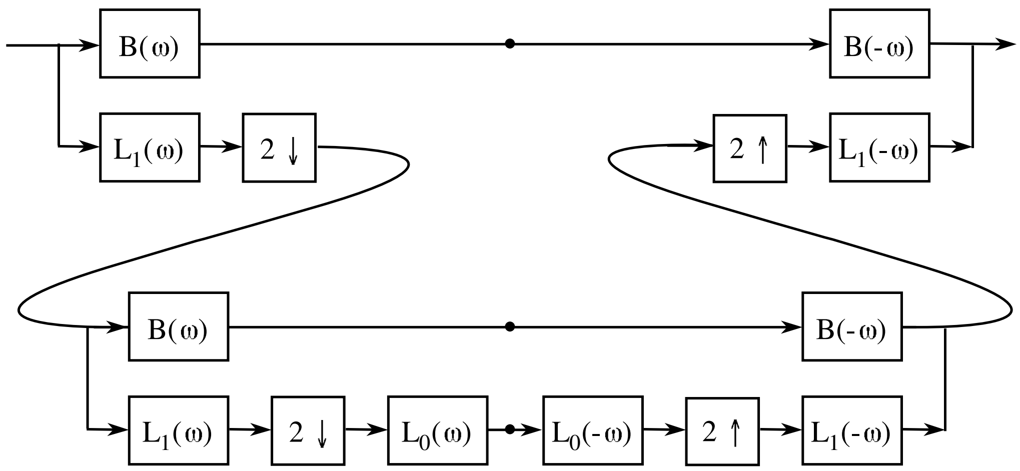


Figure 11: Illustration of a two-stage recursive cascade of the system shown in figure 10.

Since $B(\omega)$ is bandpass, and the $L_i(\omega)$ are lowpass, $S(\omega)$ has a lowpass characteristic. Therefore, a high-pass residue band must be computed by convolving with the filter:

$$R(\omega) = 1 - S(\omega)$$

Alternatively, the original image can be initially upsampled to eliminate all frequency components that would be passed by $R(\omega)$.

In order to cascade the system recursively, we must be able to replace some portion of the diagram with the entire system. We therefore require that $S(\omega) = |L_0(\omega)|^2$, thus allowing the diagram to be recursively cascaded as illustrated in figure 11. The resulting constraint on the bandpass filter $B(\omega)$ is

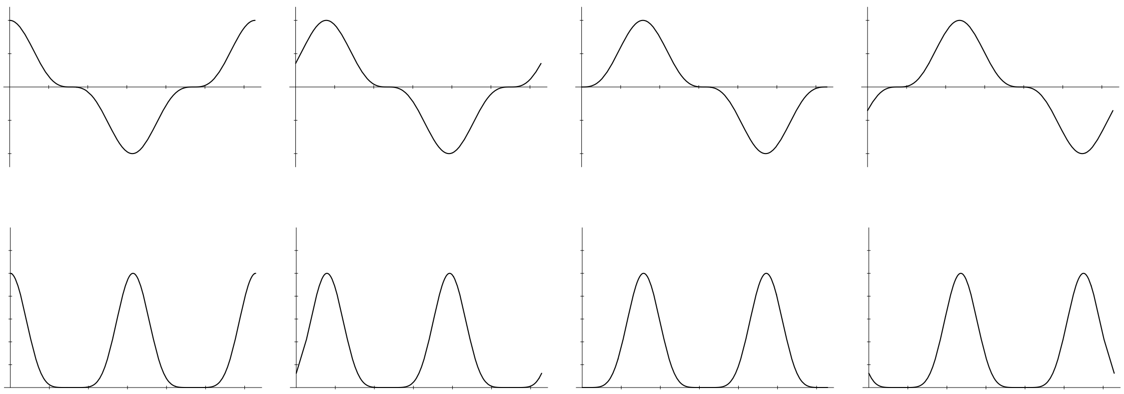
$$|L_0(\omega)|^2 = |B(\omega)|^2 + |L_1(\omega)|^2 \, |L_0(2\omega)|^2 \tag{14}$$

This recursion constraint is used in the design of the radial filters.

One other constraint on the radial filter design is that the subsampling operation should not introduce significant aliasing in the lowpass branch. It would seem that this constrains the lowpass filter $L_1(\omega)$ to have strictly zero response above $\omega = \pi/2$. In practice, the restriction is less severe. The lowpass filter $L_0(\omega)$ that follows the subsampling operation removes most aliased components, which are high frequency in the subsampled domain. Therefore, we used a seven-tap binomial lowpass filter that is fairly gentle in the frequency domain: $l_1[n] = \frac{1}{64} \cdot [1, 6, 15, 20, 15, 6, 1]$.

We also have freedom to choose the $L_0(\omega)$ filter, or equivalently, the system response $S(\omega)$. Since it represents the lowpass response of the overall system, we require that it be unity from 0 to $\frac{\pi}{2}$ radians, and zero at $\omega = \pi$. Using the Parks-McClellan algorithm, we constructed the 13-tap filter that best meets these criteria.

Having specified the two lowpass filters, the bandpass filter is constrained by the recursion relation given in equation (14). We designed a symmetric 15-tap bandpass filter that minimizes the maximum error amplitude. We used a simplex algorithm [38] to search the eight-dimensional space of free parameters. The result is the bandpass filter response shown in figure 13(a). The maximum power deviation from the desired frequency response is roughly 3.5 percent.

(a) 

(b)

Figure 12: (a) Set of four basis functions of the form $\cos^3(\theta - n\Delta_\theta)$, plotted from 0 to 2π. This set of basis functions is sufficient to synthesize $\cos^3(\theta - \theta_0)$, for any θ_0. (b) Basis function power as a function of angle.

Angular Frequency Component Design

We chose an angular frequency reponse $H(\theta) = j\cos^3(\theta)$. This can be expressed in terms of sinusoidal harmonics through use of a standard trigonometric identity:

$$\cos^3(\theta) = \tfrac{1}{4}\cos(3\theta) + \tfrac{3}{4}\cos(\theta).$$

Then the number of angular basis functions required for shiftability and the interpolation functions are determined by equation (5). Solving for this case gives the exact interpolation functions,

$$b_n(\theta) = \tfrac{1}{4}\left[2\cos(\theta - n\Delta_\theta) + 2\cos(3(\theta - n\Delta_\theta))\right] \tag{15}$$

where $\Delta_\theta = \frac{\pi}{4}$, $n \in \{0, 1, 2, 3\}$.

One-dimensional plots of the angular Fourier components of the four angular basis functions, and their power spectra, are shown in figure 12. A linear combination of the four basis functions in figure 12(a) can synthesize an arbitrary angular translation of the $\cos^3(\theta)$ kernel.

For some applications, we need to employ quadrature pairs of filters that have the same frequency response but differ in phase by $\frac{\pi}{2}$ radians (i.e. pairs of filters that are Hilbert transforms of each other [4]). Since the $\cos^3(\theta)$ frequency response is odd-symmetric, its Hilbert transform is $|\cos^3(\theta)|$. To make a steerable set of the quadrature complements of the bandpass filters, we use the first three terms in a Fourier expansion of the Hilbert transform:

$$\left|\cos^3(\theta)\right| \approx 0.4244 + 0.5093\cos(2\theta) + 0.0727\cos(4\theta). \tag{16}$$

This angular function requires five basis functions for shiftability. We note that since the power spectra of both filters of the quadrature pair are identical, either set of filters can be used in the implementation of the pyramid.

Two-dimensional Filter Design

Using the frequency transformation method [26], we converted the one-dimensional radial filters into two-dimensional filters. Figure 13(b) shows the two-dimensional bandpass filter.

The frequency transformation method produces two-dimensional filters with only approximate circular symmetry, but the approximation is quite accurate over the passband of our filters.

To construct a set of oriented filters, we subdivide the annular bandpass spectrum into orientation subbands. The angular variation is sufficiently slow to use the frequency sampling method [33]. We computed the Fourier transform of the bandpass kernel, multiplied by the four desired angular responses, $\cos^3(\theta - n\Delta_\theta)$, and computed the inverse Fourier transform to obtain the basis filter impulse responses. Figures 13(c)–(f) show the frequency responses of the bandpass filters.

We can now see that the steerable pyramid transform is (approximately) self-inverting over the lowpass range of radial frequencies $\omega_r \in [0, \frac{\pi}{2}]$. When applying the oriented filters in two dimensions, the convolution results are not spatially subsampled. These subbands are therefore spatially shiftable. The lowpass radial filters were designed to prevent aliasing, and so the subsampled lowpass signal is also spatially shiftable. In the frequency domain, both the angular and radial component designs ensure that the sum of squares of the basis functions is constant with respect to orientation and over the relavant range of scales (i.e., over the passband of $L_0(\omega)$). Therefore, by proposition 5, the pyramid is self-inverting.

6 Image Processing Applications

Since the steerable pyramid is jointly shiftable in orientation and position, one can perform useful image analysis and manipulation directly on the transform representation. In [15], the steerable pyramid representation was used to infer surface shape from image intensities (known as the "shape-from-shading" problem). Here, we demonstrate its use in two more applications.

Stereo Matching

In the stereo matching problem, the visual system is confronted with two views of a scene from differing positions (left and right eyes). The task is to find the relative horizontal displacement between corresponding points in the images. This displacement is inversely proportional to the distance to the point in the three-dimensional world, and so may be used to recover a "depth map" for the image. Because matching is simplest when there are small displacements between the two images, it is common to apply stereo matching algorithms within a multi-resolution representation. The matching is first performed at coarse spatial scales, where all displacements are small relative to the distance between pixels, and the results are then used as initial estimates for progressively finer spatial scales. Thus, at each stage of the computation, the matcher operates on measurements at the current scale, and estimates from the previous scales.

Since the steerable pyramid is a multi-scale representation, it may be used to implement a coarse-to-fine disparity estimator. More importantly, the spatial shiftability of the subbands allows accurate displacement estimation directly from the subband coefficients. Without the spatial shiftability property, a uniform translation of both images would transfer coefficient energy into other orientation and frequency (scale) bands. Thus a disparity estimator operating on a subband would be affected by the absolute spatial positions of the images.

Figure 14 shows an example. Figures 14(a) shows the left and right eye views of a random

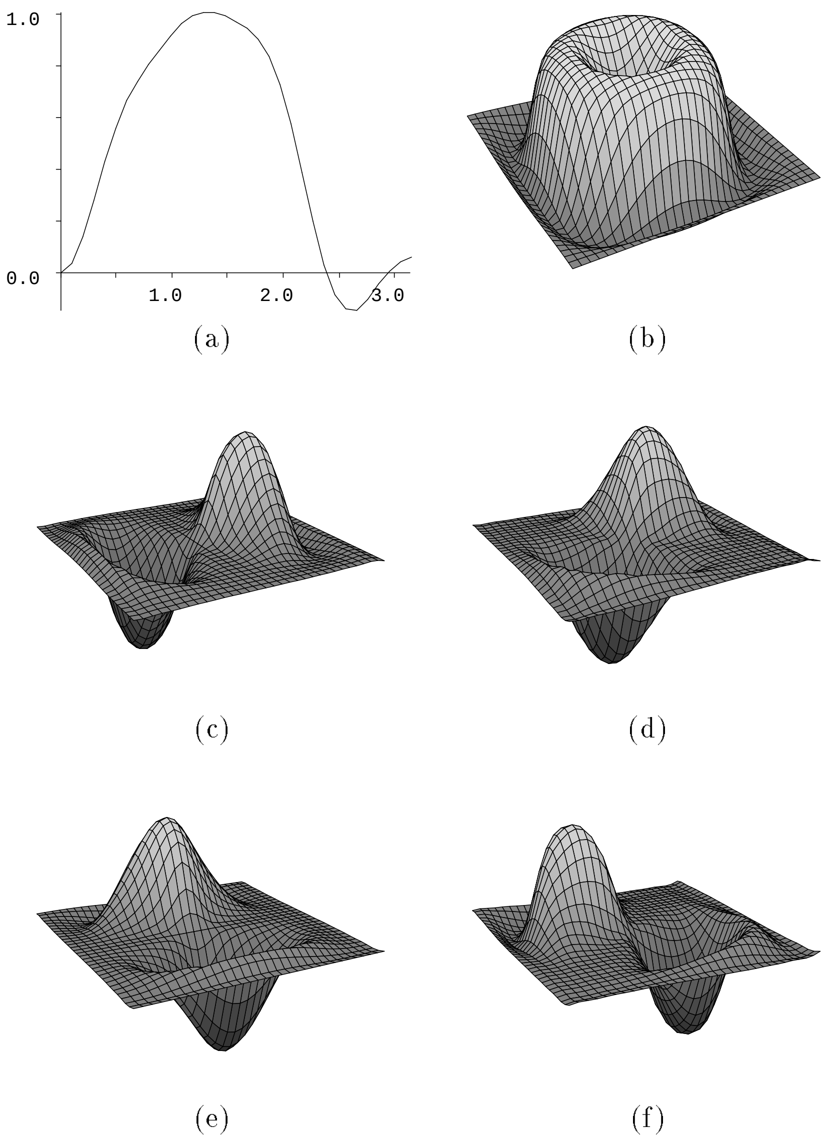


Figure 13: Frequency domain filter response plots, illustrating design procedure for digital steerable filter. (a) Desired radial frequency distribution, plotted from 0 to π. (b) Desired angularly symmetric two-dimensional frequency response, obtained through frequency transformation. The profile in (b) was multiplied by the desired $\cos^3(\theta - n\theta)$ angular frequency responses and inverse transformed to yield the steerable basis set. (c) – (f) The imaginary component of the frequency responses of the resulting steerable filters.

dot stereogram [21]. A central rectangle of dots is displaced horizontally by one pixel between the left and right eye views. Figure 14(b) shows an overlaid plot of cross-sections of the same horizontal line of the two images; (c) shows a surface plot of the depth map corresponding to this stereo pair.

We used this test stimulus to compare the performance of the steerable pyramid and a wavelet image representation in stereo matching. We compare the disparity estimates after one stage in a coarse-to-fine algorithm (i.e., we compute stereo disparity at a single spatial scale).

We use a least-squares gradient-based displacement estimator similar to that used by Lucas and Kanade [27]. This is based on a locally planar model of the image. For each patch P of the image, it computes the image displacement that minimizes the squared error function:

$$E(d) = \sum_{x \in P} \left[I_r(x + \tfrac{d}{2}) - I_l(x - \tfrac{d}{2}) \right]^2 ,$$

where $I_l(x)$ and $I_r(x)$ are the left and right images, and d is the disparity at the center of patch P. Approximating the displaced images with a first-order Taylor expansion and solving for the minimum-error estimate gives:

$$\hat{d} = \frac{\displaystyle\sum_{x \in P} \frac{\partial \bar{I}(x)}{\partial x} \left[I_r(x) - I_l(x) \right]}{\displaystyle\sum_{x \in P} \left[\frac{\partial \bar{I}(x)}{\partial x} \right]^2}$$

where $\bar{I}(x)$ is the pointwise average of the two images. In our example, we used a patch size of nine pixels, weighted by a gaussian profile.

The steerable pyramid coefficients corresponding to a single orientation and scale, are shown in figure 14(g). Figure 14(h) shows graphs of the left half of the corresponding horizontal cross-sections. The left portions of the cross-sections are identical. Since the subbands are not aliased, the right portions of the two cross-sections are approximately translations of each other (if we had interpolated them using the correct interpolator, they would be exact translations). Use of the Lucas and Kanade algorithm on these subbands produces a good estimate of disparity. The resulting depth map is shown in figure 14(i).

Figure 14(d) shows the Daubechies 4-tap wavelet representation of the stereo pair reconstructed from a single subband chosen to match that used in the steerable pyramid example. Figure 14(e) shows the left half of the corresponding plots of horizontal cross-sections. Note that the right portions of the two cross-sections are *not* related by translation. As in the example of figure 1, translation of the input signal causes an exchange of power amongst the subbands, and this aliasing in the representation produces large errors in the disparity estimates. The resulting depth estimate is poor, as can be seen in figure 14(f).

We emphasize that the choice of a different displacement estimation technique, such as correlation, will not solve this problem: the aliasing in the subbands causes an irreversible loss of information. One could apply a lowpass filter to each subband to eliminate those frequencies that are contaminated by aliasing, but this would leave gaps in the spectrum which would be ignored by the estimator. An input signal with significant spectral content in these gaps (and not elsewhere) could not be processed.

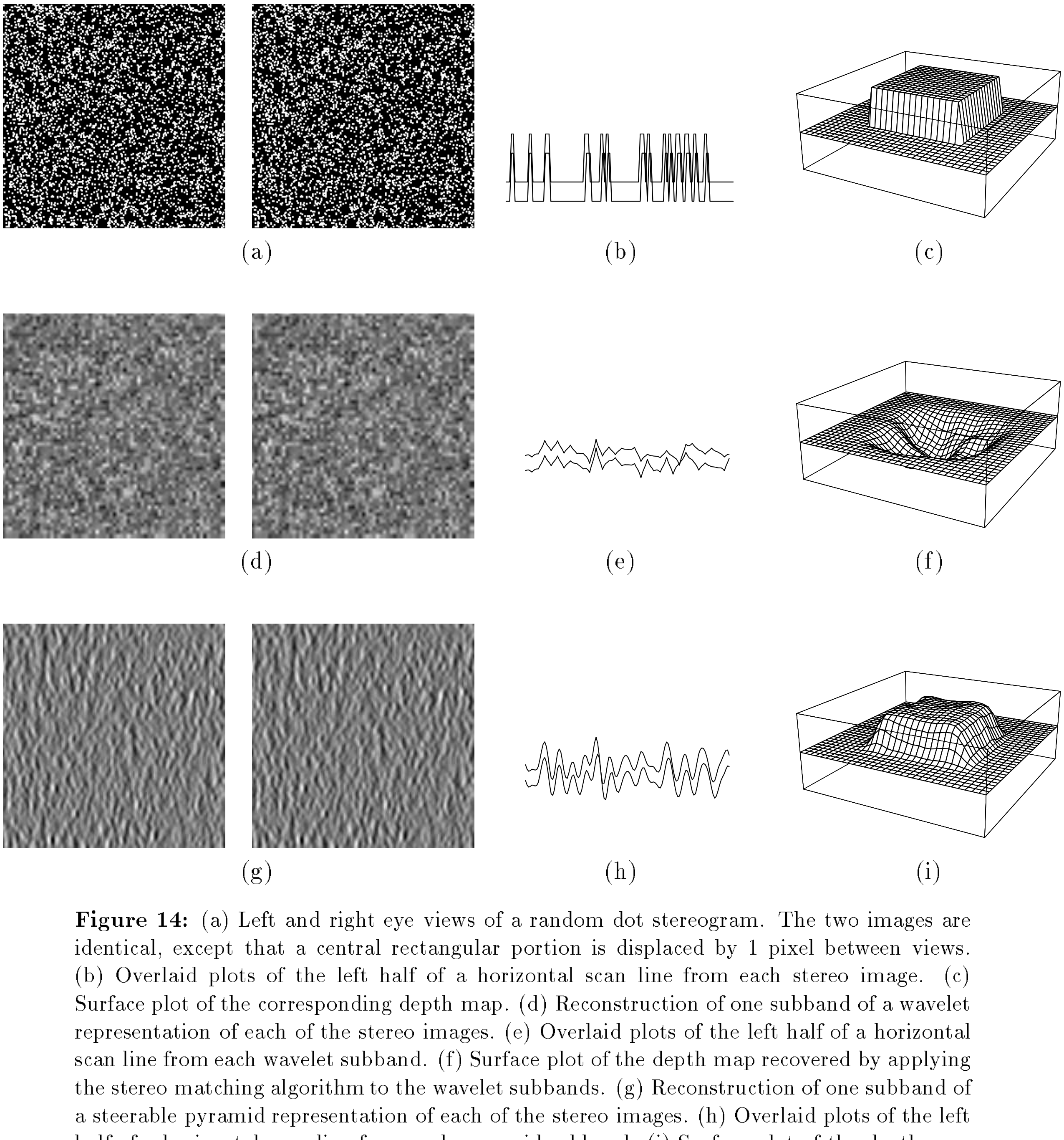


Figure 14: (a) Left and right eye views of a random dot stereogram. The two images are identical, except that a central rectangular portion is displaced by 1 pixel between views. (b) Overlaid plots of the left half of a horizontal scan line from each stereo image. (c) Surface plot of the corresponding depth map. (d) Reconstruction of one subband of a wavelet representation of each of the stereo images. (e) Overlaid plots of the left half of a horizontal scan line from each wavelet subband. (f) Surface plot of the depth map recovered by applying the stereo matching algorithm to the wavelet subbands. (g) Reconstruction of one subband of a steerable pyramid representation of each of the stereo images. (h) Overlaid plots of the left half of a horizontal scan line from each pyramid subband. (i) Surface plot of the depth map recovered by applying the stereo matching algorithm to the steerable pyramid subbands.

Image Enhancement

A common problem in image processing is to restore an image that has been corrupted by noise. Through the use of an image transform, one can decompose the image into a set of coefficients that allow discrimination between image information and noise. These coefficients can be modified according to the likelihood that they represent signal rather than noise. The "cleaned" image can then be constructed by inverting the transform. (see, for example, [37, 23, 6]).

Natural images tend to contain locally oriented structures such as lines, edges, and textures, while noise tends to be isotropic (i.e., without preferred orientation). Therefore an image decomposition based on oriented filters should help one to determine whether a given image structure is due to image or noise, and should allow one to enhance the image quality (cf. [23]). For these reasons, the steerable pyramid representation described in the previous section is well-suited for image enhancement. Shiftability in orientation and space allow all orientations and positions to be treated uniformly. The multi-scale nature of the representation allows the processing to depend on the spatial frequency band.

We want to preserve the contents of the image representation that correspond to visually significant oriented structures in the image. These structures may appear at any relative phase: even-phase (lines), odd-phase (edges), or contours of intermediate phase. A local "oriented energy" measure, formed by the sum of squares of a quadrature pair of even- and odd-phase filters, will treat structures of different phases uniformly (cf. [36]). Thus, we construct two steerable pyramids, one using odd-symmetric filters, and the other using their even-symmetric Hilbert transform counterparts, and use the sum of squares of the coefficients at each position, orientation and scale to measure orientation strength.

We use this oriented energy measure to determine the appropriate modification of the coefficients in one of the pyramids (although we use both of the quadrature pair of filters to compute energy, only one is needed to reconstruct the enhanced image). If the oriented energy is large, the pyramid coefficient at that orientation is left unchanged; otherwise it is be attenuated. To perform this operation, we use a soft threshold function similar to that described in [6]:

$$\hat{c}_i = \frac{c_i}{(1 + \exp(-S(\sqrt{e_i} - T)))}, \tag{17}$$

where c_i is the coefficient at a particular orientation, e_i is the energy at that orientation, and S and T are sharpness and threshold parameters that are chosen for each subband.

We use the steerability of the image representation to ensure that the noise processing is independent of image orientation. At each scale and position, we analytically find the local dominant orientation from the quadrature pair of steerable filters [15]. For the case of a single oriented structure this is the orientation which maximizes the energy measure e_i. We then steer each of the odd phase steerable pyramid filters so that one of them is aligned with the dominant orientation. We modify the steered coefficients according to equation 17. After this modification, we steer the coefficients back to their original orientations, and reconstruct the image from these altered coefficients. Altering the coefficients in a coordinate system defined by the local dominant orientation ensures that the image will be processed in a rotation-invariant manner.

Figure 15 (a) shows the original noise-free image, along with an enlarged portion. Pixel intensity values are in the range $[30, 224]$. Figure 15(b) shows the image corrupted by white noise uniformly distributed on the interval $[-20, 20]$. The SNR for this image is 12.42 dB, or 13.6 dB peak-to-peak. To compute a noise-reduced image, we constructed two levels of a steerable pyramid from the noisy image, and modified the coefficients as described above, using a sharpness parameter, $S = 0.5$, and a threshold, $T = 2$ for the lower frequency band and $T = 11$ for the higher frequency band. The same parameters were used for all orientations. These values were roughly chosen to minimize the mean squared error. The pyramid transform was then inverted to produce the processed image, shown in figure 15 (c). Very little noise remains in the processed image, yet most important image features are preserved. The SNR of this result is 23.0 dB.

For comparison with the steerable pyramid algorithm, we also restored the image using a Wiener filter. Implementation of the Wiener filter requires estimates of the power spectra of both the image and the noise. For the image spectrum estimate we used the average of the power spectra of an ensemble of 8 different face images, taken under conditions identical to those used for figure 15 (a). We used a flat power spectrum model for the noise, with amplitude equal to the mean of the spectral power of the actual noise. Figure 15 (d) shows the Wiener filter restoration result, with a SNR of 19.24 dB. While this value is only moderately lower than that of the steerable pyramid method, the visual appearance of the restored image is both noisier and more blurred than the steerable pyramid result.

7 Conclusion

The recent development of wavelet transforms has generated enthusiasm and controversy in a broad range of disciplines. The defining property of these transforms is that the basis functions are dilations and translations of a common kernel (in two dimensions, we add rotations to this list). Two secondary properties of typical wavelet transforms are that they are orthogonal, and they may be implemented using recursive pyramid structures. Their representational and computational structure makes them suitable for efficient signal and image coding. We have argued, however, that the aliasing that results from the critical sampling constraint (as required by orthogonality) is often problematic for applications in signal or image analysis.

In the spatial domain, one would like an image representation to treat its input in a uniform manner, regardless of the relative alignment of the input and the transform sampling lattices. No subsampled subband transform can be translation invariant, in the strong sense that the transformation operator commutes with the translation operator. But it is possible to generate transforms in which the information (and the power) contained within a given subband is invariant to translations of the input signal. Such transforms must generally be oversampled; the Nyquist criterion specifies the necessary sampling rate.

We have developed analogous concepts in the context of orientation analysis. It is possible to devise "steerable" image representations in which the power in the coefficients corresponding to different orientations (at the same scale and position) is invariant to rotations of the input signal. The concept of steerability has been explored in prior work; here we have demonstrated that it is essentially equivalent to the translation invariance issue of the spatial domain.

We also discussed transforms that are "scalable," in that measurements at any scale can be derived from those at a discrete set of scales. The scalable representations require that we take our space-domain concepts and re-state them in the log frequency domain.

We have formalized a generalization of these properties which we call "shiftability". We

(a)
(b)
(c)
(d)

Figure 15: Noise removal example. Figures on the right are enlarged portions of those on the left. (a) The original noise-free image. (b) The image corrupted by noise. SNR is 12.42 dB. (c) Results of image restoration using steerable pyramid. SNR is 23.0 dB (d) Results of image restoration using a Wiener filter (see text). SNR is 19.24 dB.

showed that shiftability is equivalent to the invariance of the transform power to input signal translation (proposition 2), and that a sub-property of shiftability is that the sum of squares of the basis functions is constant (corollary 3). We discussed the property of self-invertibility and showed that this property holds for transforms that are shiftable in space and have flat basis function power in the frequency domain (proposition 5). We also discussed the issue of joint shiftability.

We have demonstrated these ideas by designing one transform that is jointly shiftable in space and orientation (a "steerable pyramid"), and another that is jointly shiftable in space and scale. Although these transforms are less efficient than critically sampled representations, we feel that for many applications, the benefits of shiftability are worth the cost. We demonstrate advantages of the transforms by applying them to the problems of stereo matching, scale-space analysis, and image enhancement.

There are open issues remaining. The relationship of the shiftability property to the mathematics of groups could be explored. A measure of joint shiftability might be used in place of Gabor's joint localization constraint to develop approximately jointly shiftable transforms for a given sampling structure. The practical problems in the design of scalable transforms should be investigated further. In particular, it is of interest to construct a scalable transform that is efficiently implemented as a recursive pyramid. Finally, the shiftability concept could be extended to other domains such as time and color.

Acknowledgements

We thank John Wang for his help in implementing the stereo matcher. We also thank Rosalind W. Picard, Stephen Scherock, and the reviewers of this paper for their insightful comments and suggestions. Support for this work was provided by contracts with Goldstar Co, Ltd., the Television of Tomorrow Program at the MIT Media Laboratory, NASA RTOP 506-71-51, and DARPA/RADC F30602-89-C-0022.

References

[1] E H Adelson, C H Anderson, J R Bergen, P J Burt, and J M Ogden. Pyramid methods in image processing. *RCA Engineer*, 29(6):33–41, November/December 1984.

[2] E H Adelson and E P Simoncelli. Truncated subband coding of images. U.S. Patent #4,917,812, 1989.

[3] E H Adelson, E P Simoncelli, and R Hingorani. Orthogonal pyramid transforms for image coding. In *Proc. SPIE*, volume 845, pages 50–58, Cambridge, MA, October 1987.

[4] R N Bracewell. *The Fourier Transform and its Applications.* McGraw-Hill, 1978.

[5] P J Burt and E H Adelson. The Laplacian pyramid as a compact image code. *IEEE Trans. Comm.*, COM-31(4):532–540, April 1983.

[6] C R Carlson, E H Adelson, and C H Anderson. System for coring an image-representing signal. U.S. Patent #4,523,230, 1985.

[7] A Croisier, D Esteban, and C Galand. Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques. In *Intl. Conf. on Information Sciences and Systems*, pages 443–446, Patras, August 1976.

[8] I Daubechies. Frames of coherent states – phase space localization and signal analysis. Unpublished manuscript.

[9] I Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, 41:909–996, 1988.

[10] I Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Info. Theory*, IT-36(5), September 1990.

[11] J G Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A*, 2(7):1160–1169, July 1985.

[12] J G Daugman. Complete discrete 2–D Gabor transforms by neural networks for image analysis and compression. *IEEE Trans. Acoust. Speech Signal Proc.*, 36(7):1169–1179, 1988.

[13] D Esteban and C Galand. Application of quadrature mirror filters to split band voice coding schemes. In *Proc. ICASSP*, pages 191–195, 1977.

[14] W T Freeman and E H Adelson. Steerable filters. In *Topical Meeting on Image Understanding and Machine Vision*, volume 14. Optical Society of America, June 1989. Technical Digest Series.

[15] W T Freeman and E H Adelson. The design and use of steerable filters. *IEEE Pat. Anal. Mach. Intell.*, 13(9):891–906, 1991.

[16] D Gabor. Theory of communication. *J. IEE*, 93:492–457, 1946.

[17] H Gharavi and A Tabatabai. Sub-band coding of digital images using two-dimensional quadrature mirror filtering. In *Proc. of SPIE*, volume 707, pages 51–61, 1986.

[18] R A Gopinath and C S Burrus. Efficient computation of the wavelet transforms. In *ICASSP*, pages 1599–1601, 1990.

[19] G H Granlund. In search of a general picture processing operator. *Comp. Graphics Image Proc.*, 8:155–173, 1978.

[20] A Grossmann and J Morlet. Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Math.*, 15:723–736, 1984.

[21] B Julesz. Experiments in the visual perception of texture. *Scientific American*, 232(4):34–43, 1975.

[22] H Knutsson and G H Granlund. Texture analysis using two-dimensional quadrature filters. In *IEEE Computer Society Workshop on Computer Architecture for Pattern Analysis and Image Database Management*, pages 206–213, 1983.

[23] H. Knutsson, R. Wilson, and G. H. Granlund. Anisotropic nonstationary image estimation and its applications: Part 1 – Restoration of noisy images. *IEEE Trans. Comm.*, 31(3):388–397, 1983.

[24] J J Koenderink. The structure of images. *Biol. Cybern.*, 50:363–370, 1984.

[25] R M Lerner. *Lectures on Communication System Theory*, chapter 10. McGraw-Hill, New York, 1961.

[26] J Lim. *Two-Dimensional Signal and Image Processing*. Prentice Hall, Englewood Cliffs, New Jersey, 1990.

[27] B D Lucas and T Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. Seventh IJCAI*, pages 674–679, Vancouver, 1981.

[28] S G Mallat. Multifrequency channel decompositions of images and wavelet models. *IEEE Trans. Acoust. Speech Signal Proc.*, 37(12):2091–2110, 1989.

[29] S G Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Pat. Anal. Mach. Intell.*, 11:674–693, July 1989.

[30] S G Mallat. Zero-crossings of a wavelet transform. *IEEE Trans. Info. Theory*, 37(4):1019–1033, July 1991.

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

[32] J B Martens. The Hermite transform – theory. *IEEE Trans. Acoust. Speech Signal Proc.*, 38(9):1595–1606, 1990.

[33] A V Oppenheim and R W Schafer. *Digital Signal Processing.* Prentice-Hall, Englewood Cliffs, NJ, 1975.

[34] A P Pentland. Fractal based description of natural scenes. *IEEE Pat. Anal. Mach. Intell.*, 6(6):661–674, 1984.

[35] P Perona. Deformable kernels for early vision. In *IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition*, pages 222–227, Maui, 1991.

[36] P Perona and J Malik. Detecting and localizing edges composed of steps, peaks and roofs. In *Proc. 3rd Intl. Conf. Computer Vision*, Osaka, Japan, 1990.

[37] P G Powell and B E Bayer. A method for the digital enhancement of unsharp, grainy photographic images. In *IEE Intl. Conf. Electronic Image Processing*, number 214, page 179, 1982.

[38] W H Press, B P Flannery, S A Teukolsky, and W T Vetterling. *Numerical Recipes in C.* Cambridge University Press, New York, 1988.

[39] E P Simoncelli. Orthogonal sub-band image transforms. Master's thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA, May 1988.

[40] E P Simoncelli and E H Adelson. Non-separable extensions of quadrature mirror filters to multiple dimensions. In *Proc. IEEE: Special Issue on Multi-dimensional Signal Processing*, April 1990.

[41] E P Simoncelli and E H Adelson. Subband transforms. In John W. Woods, editor, *Subband Image Coding*, chapter 4. Kluwer Academic Publishers, Norwell, MA, 1990.

[42] G T Strang. Wavelets and dilation equations: A brief introduction. *SIAM Review*, 31(4):614–627, December 1989.

[43] A Tran, K Liu, K Tzou, and E Vogel. An efficient pyramid image coding system. In *Proc. ICASSP*, volume 2, pages 18.6.1–18.6.4, 1987.

[44] P P Vaidyanathan. On power-complementary FIR filters. *IEEE Trans. Circuits and Systems*, CAS-32(12):1308–1310, December 1985.

[45] M Vetterli. Multidimensional subband coding: some theory and algorithms. *Signal Processing*, 6(2):97–112, 1984.

[46] M Vetterli. A theory of multirate filter banks. *IEEE Trans. Acoust. Speech Signal Proc.*, ASSP-35(3):356–372, March 1987.

[47] A B Watson. The cortex transform: rapid computation of simulated neural images. *Comp. Vis. Graphics Image Proc.*, 39:311–327, 1987a.

[48] A P Witkin. Scale-space filtering. In *Proc. Intl. Joint Conf. Artificial Intelligence*, pages 1019–1021, 1985.

[49] J W Woods, editor. *Subband Image Coding.* Kluwer Academic Publishers, Norwell, MA, 1990.

[50] J W Woods and S D O'Neil. Subband coding of images. *IEEE Trans. Acoust. Speech Signal Proc.*, ASSP-34(5):1278–1288, October 1986.