Presented at: 2nd IEEE International Conference on Image Processing.
Washington, DC. vol III, pp 444-447. October, 1995.

THE STEERABLE PYRAMID: A FLEXIBLE ARCHITECTURE FOR MULTI-SCALE DERIVATIVE COMPUTATION

Eero P Simoncelli
GRASP Laboratory, Room 335C
3401 Walnut St
Philadelphia, PA 19104-6228

William T Freeman
Mitsubishi Electric Research Laboratories
201 Broadway
Cambridge, MA 02139

ABSTRACT

We describe an architecture for efficient and accurate linear decomposition of an image into scale and orientation subbands. The basis functions of this decomposition are directional derivative operators of any desired order. We describe the construction and implementation of the transform. [1]

Differential algorithms are used in a wide variety of image processing problems. For example, gradient measurements are used as a first stage of many edge detection, depth-from-stereo, and optical flow algorithms. Higher-order derivatives have also been found useful in these applications. Extraction of these derivative quantities may be viewed as a decomposition of a signal via terms of a local Taylor series expansions [1].

Another widespread tool in signal and image processing is multi-scale decomposition. Apart from the advantages of decomposing signals into information at different scales, the typical recursive form of these algorithms leads to large improvements in computational efficiency.

Many authors have combined multi-scale decompositions with differential measurements (eg., [2, 3]). In these cases, a multi-scale pyramid is constructed, and then differential operators (typically, differences of neighboring pixels) are applied to the subbands of the pyramid. Since both the pyramid decomposition and the derivative operation are linear and shift-invariant, we may combine them into a single operation. The advantages of doing so are that the resulting derivatives may be more accurate (see [4]). In this paper, we propose a simple, efficient decomposition architecture for combining these two operations.

The decomposition is the latest incarnation of "steerable pyramid", as developed in [5, 6]. Similar representations have been developed by Perona and collaborators [7]. In this linear decomposition, an image is subdivided into a collection of subbands localized in both scale and orientation. The scale tuning of the filters is constrained by a recursive system diagram (described below). The orientation tuning is constrained by the property of steerability [5]. A set of filters form a steerable basis if (1) they are rotated copies of each other, and (2) a copy of the filter at any orientation may be computed as a *linear combination* of the basis filters. The simplest example of a steerable basis is a set of $N+1$ Nth-order directional derivatives.

In addition to having steerable orientation subbands, the transform we describe is designed to be "self-inverting" (i.e., the matrix corresponding to the inverse transformation is equal to the transpose of the forward transformation matrix)[2], and is essentially aliasing-free. Most importantly, the pyramid can be designed to produce any number of orientation bands, k. The resulting transform will be overcomplete by a factor of $4k/3$.

A summary of these properties, in comparison with two well-known multi-scale decompositions is given in table 1. Note that the steerable pyramid retains some of the advantages of orthonormal wavelet transforms (eg., basis functions are localized in space and spatial-frequency; the transform is a tight frame), but improves on some of their disadvantages (eg., aliasing is eliminated; steerable orientation decomposition). One obvious disadvantage is in computational efficiency: the steerable pyramid is substantially overcomplete.

We now describe the steerable pyramid in more detail. The decomposition is most easily defined in the Fourier domain, where it is (ideally) polar-

[1]Source code and filter kernels for implementation of the steerable pyramid are available via anonymous ftp from `ftp.cis.upenn.edu:pub/eero/steerpyr.tar.Z`

[2]In the wavelet literature, such a transform is known as a *tight frame* [8]

	Laplacian Pyramid	Dyadic QMF/Wavelet	Steerable Pyramid
self-inverting (tight frame)	no	yes	yes
overcompleteness	$4/3$	1	$4k/3$
aliasing in subbands	perhaps	yes	no
rotated orientation bands	no	only on hex lattice [9]	yes

Table 1: Properties of the Steerable Pyramid relative to two other well-known multi-scale representations.

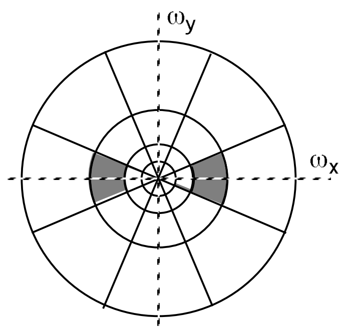

Figure 1. Idealized illustration of the spectral decomposition performed by a steerable pyramid with $k = 4$. Frequency axes range from $-\pi$ to π. The basis functions are related by translations, dilations and *rotations* (except for the initial highpass subband and the final lowpass subband). For example, the shaded region corresponds to the spectral support of a single (vertically-oriented) subband.

separable. Figure 1 contains a diagram of the idealized frequency response of the subbands, for $k = 4$. We write the Fourier magnitude of the ith oriented bandpass filter in polar-separable form:

$$B_i(\vec{\omega}) = A(\theta - \theta_i)B(\omega),$$

where $\theta = \tan^{-1}(\omega_y/\omega_x)$, $\theta_i = \frac{2\pi}{k}$ and $\omega = |\vec{\omega}|$. Below, we describe the constraints on the two components $A(\theta)$ and $B(\omega)$.

1. ANGULAR DECOMPOSITION

The angular portion of the decomposition, $A(\theta)$, is determined by the desired derivative order. A directional derivative operation in the spatial domain corresponds to multiplication by a linear ramp function in the Fourier domain, which we rewrite in polar coordinates as follows:

$$-j\omega_x = -j\omega\cos(\theta)$$

(note that we have described a derivative operator in the x direction). We ignore the imaginary constant, and the factor of ω, which is absorbed into the radial portion of the function. The relevant angular portion of the first derivative operator (in the x direction) is thus $\cos(\theta)$.

Higher-order directional derivatives correspond to multiplication in the Fourier domain by the ramp raised to a power, and thus the angular portion of the filter is $\cos(\theta)^N$ for an Nth-order directional derivative. Knuttson and Granlund have also developed polar-separable filters with such angular components [10]. The steerability of such functions has been discussed in our previous work [5, 6].

2. RADIAL DECOMPOSITION

The radial function, $B(\omega)$, is constrained by both the desire to build the decomposition recursively (i.e., using a "pyramid" algorithm), and the need to prevent aliasing from occurring during subsampling operations. The recursive system diagram for $B(\omega)$ is given in figure 2[3].

The filters $H_0(\omega)$ and $L_0(\omega)$ are necessary for preprocessing the image in preparation for the recursion. The recursive portion of the diagram corresponds to the subsystem contained in the dashed box. This subsystem decomposes a signal into two portions (lowpass and highpass). The lowpass portion is subsampled, and the recursion is performed by repeatedly applying the recursive transformation to the lowpass signal.

The constraints on the filters in the diagram are as follows:

1. Bandlimiting (to prevent aliasing in the subsampling operation):

$$L_1(\omega) = 0 \qquad \text{for} |\omega| > \pi/2.$$

2. Flat System Response:

$$|H_0(\omega)|^2 + |L_0(\omega)|^2 \left[|L_1(\omega)|^2 + |B(\omega)|^2\right] = 1.$$

3. Recursion:

$$|L_1(\omega/2)|^2 = |L_1(\omega/2)|^2 \left[|L_1(\omega)|^2 + |B(\omega)|^2\right].$$

Typically, we choose $L_0(\omega) = L_1(\omega/2)$, so that the initial lowpass shape is the same as that used within the recursion. An idealized illustration of filters that satisfy these constraints is given in figure 3. Note that $L_1(\omega)$ is strictly bandlimited, and $B(\omega)$ is power-complementary.

[3]This system diagram is modified from that of [6].

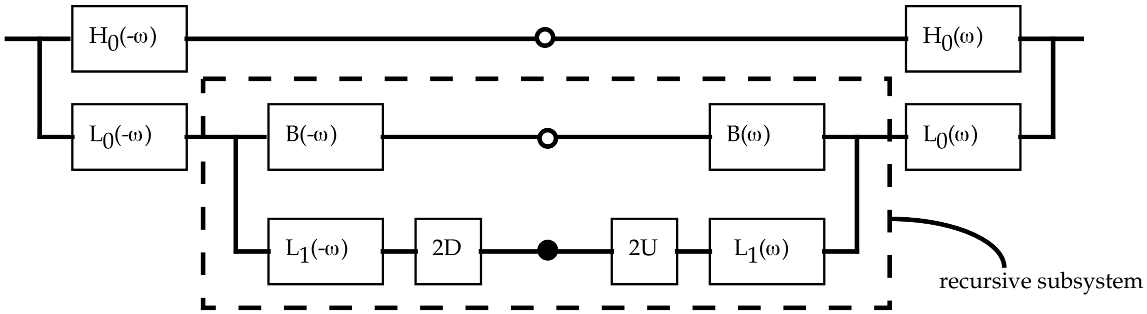

Figure 2. System diagram for the radial portion of the steerable pyramid, illustrating the filtering and sampling operations, and the recursive construction. Boxes containing "2D" and "2U" correspond to downsampling and upsampling by a factor of 2. All other boxes correspond to standard 2D convolution. The circles correspond to the transform coefficients. The recursive construction of a pyramid is achieved by inserting a copy of the diagram contents enclosed by the dashed rectangle at the location of the *solid* circle (i.e., the lowpass branch).

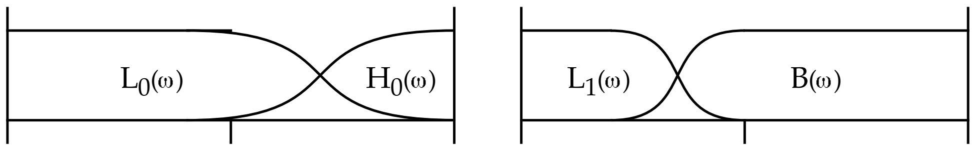

Figure 3. Idealized depiction of filters satisfying the constraints of the block diagram in figure 2. Plots show Fourier spectra over the range $[0, \pi]$.

3. IMPLEMENTATION

We have designed filters using weighted least squares techniques in the Fourier domain to approximately fit the constraints detailed above. The resulting filters are fairly compact (typically 9×9 taps) and accurate (reconstruction error on the order of 45dB). Such filters may be designed for different values of k, depending on the application. For example, a design with a single band at each scale ($k = 1$) serves as a (self-inverting) replacement for the Laplacian pyramid. A design with two bands ($k = 2$) will compute multi-scale image gradients, which may be used for computations of local orientation, stereo disparity or optical flow. Higher values of k correspond to higher order terms in a multi-scale Taylor series.

Figure 4 illustrates a 3-level steerable pyramid decomposition of a disk image, with $k = 1$. Shown are the bandpass images and the final lowpass image (the initial highpass image is *not* shown). As noted above, this pyramid may be used in applications where the Laplacian pyramid has been found useful, such as in image coding. The advantage is that the steerable pyramid is self-inverting, and thus the errors introduce by quantization of the subbands will not appear as out-of-band distortions upon reconstruction.

Figure 5 illustrates a 3-level steerable pyramid decomposition with $k = 3$. The filters are directional second derivatives oriented at $\theta \in \{-2\pi/3, 0, 2\pi/3\}$. Such a decomposition can be used for orientation analysis, edge detection, etc.

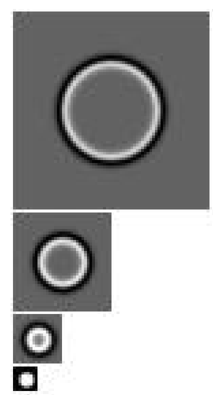

Figure 4. A 3-level $k = 1$ (non-oriented) steerable pyramid. Shown are the bandpass images and the final lowpass image.

We have explored the use of this decomposition in a number of applications, including image enhancement, orientation decomposition and junction identification, texture blending, depth-from-stereo, and optical flow. Space limitations prevent full description of these applications here; some previous results are described in [5, 6].

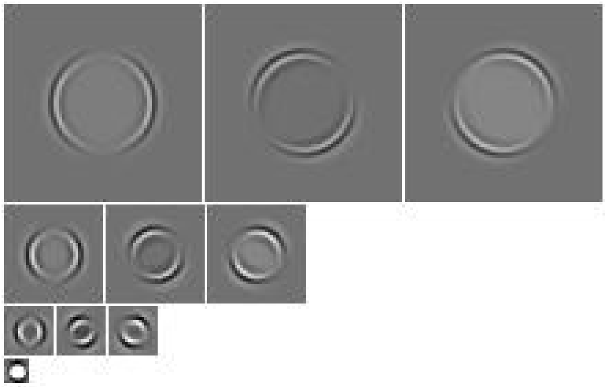

Figure 5. A 3-level $k = 3$ (second derivative) steerable pyramid. Shown are the three bandpass images at each scale and the final lowpass image.

4. REFERENCES

[1] J. J. Koenderink and A. J. van Doorn. Representation of local geometry in the visual system. *Biol. Cybern.*, 55:367–375, 1987.

[2] 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.

[3] Peter Burt. Pyramid methods in image processing. In *Proceedings of the First International Conference on Image Processing*, Austin, November 1994. plenary presentation.

[4] Eero P Simoncelli. Design of multidimensional derivative filters. In *First International Conference on Image Processing*, Austin, Texas, November 1994.

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

[6] Eero P Simoncelli, William T Freeman, Edward H Adelson, and David J Heeger. Shiftable multi-scale transforms. *IEEE Trans. Information Theory*, 38(2):587–607, March 1992. Special Issue on Wavelets.

[7] H. Greenspan, S. Belongie, R. Goodman, P. Perona, S. Rakshit, and C. H. Anderson. Overcomplete steerable pyramid filters and rotation invariance. In *Proceedings, CVPR*, pages 222–228, 1994.

[8] I Daubechies. Ten lectures on wavelets. In *Regional Conference: Society for Industrial and Applied Mathematics*, Philadelphia, PA, 1992.

[9] Eero P Simoncelli and Edward H Adelson. Non-separable extensions of quadrature mirror filters to multiple dimensions. *Proceedings of the IEEE: Special Issue on Multidimensional Signal Processing*, April 1990. Also available as MIT Media Laboratory Vision and Modeling Technical Report #119.

[10] 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.