

Simplified Design of Steerable Pyramid Filters

K. R. Castleman, M. Schulze, Q. Wu

Perceptive Scientific Instruments, Inc.
2525 South Shore Blvd.
League City, Texas 77573, USA

ABSTRACT

The steerable pyramid decomposition is an invertible representation similar to the two-dimensional discrete wavelet transform, but with interesting shift- and rotation-invariance properties. It is slightly overcomplete and amenable to a filter bank implementation via convolution and down- and up-sampling operations. This paper presents a simple method for designing the FIR filter kernels required to implement the transform.

1. INTRODUCTION

The steerable pyramid algorithm is an invertible multiscale image transform [1,2]. Figure 1 shows the iterated filter bank structure proposed by Simoncelli, et. al. for the steerable pyramid decomposition and reconstruction [1]. An outgrowth of the Laplacian pyramid [3], it decomposes an image into oriented, bandpass filtered components at different (binary) scales. It has useful shiftability properties in both translation and rotation [4].

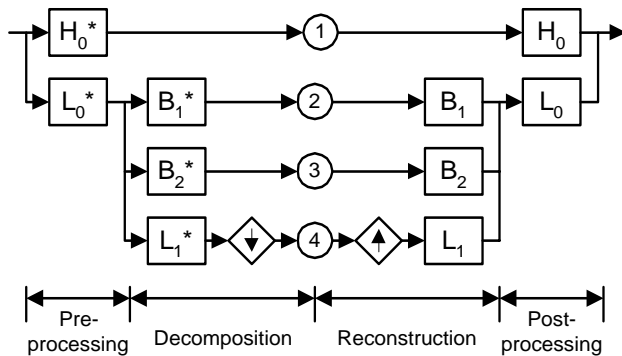


Figure 1. Block diagram of the steerable pyramid transform for $k=2$. The decomposition/reconstruction filter bank is iterated at node 4. The down- and up-arrows indicate binary down- and up-sampling, respectively.

Like the discrete wavelet transform (DWT, Mallat algorithm, [5]), the steerable pyramid transform decomposes an image into a series of scaled component images from which the original can be reconstructed. Like the DWT it subsamples the image at each (binary scaled) stage of iteration, producing progressively half-sized images. Unlike the DWT it avoids aliasing when subsampling [1]. Also unlike the DWT, it is shift- and rotationally-invariant, due to the absence of aliasing in the down-sampling process.

One must choose the order of the derivative upon which the steerable pyramid transform is based. The number, k , of

orientation bands at each scale is one more than the order of the derivative. The transform is overcomplete by the factor $4k/3$.

Each of the blocks in Figure 1 represents a two-dimensional transfer function. $H_0(u,v)$ is a highpass filter that passes the high frequency information that falls in the “corners” of the (square) frequency space through Node 1. $L_0(u,v)$ is a complementary lowpass filter, where u and v are frequency variables corresponding to the x - and y -directions, respectively. The asterisks in Fig. 1 indicate 180° rotation of the transfer function. Nodes 2 and 3 pass the high frequency (edge) information at the first scale. Node 4 passes the low-frequency information, and this is where additional stages of decomposition and reconstruction can be inserted.

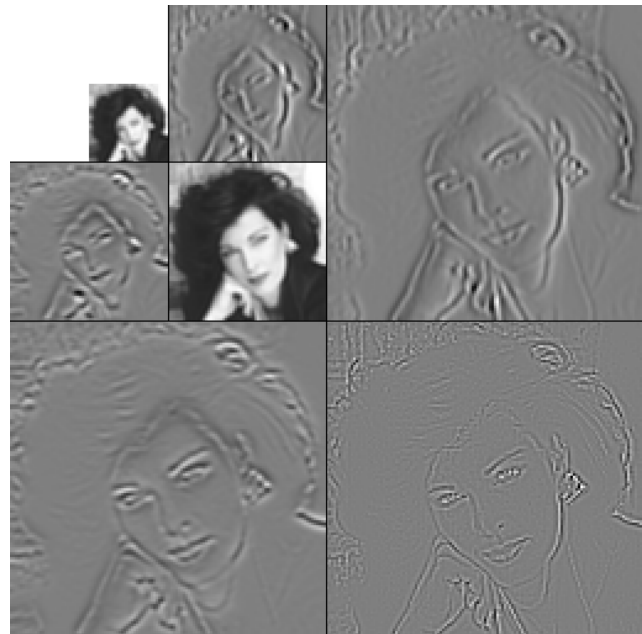


Figure 2. Two-stage steerable pyramid transform for $k=2$.

In Figure 2, the output of the (single) highpass filter fills the LR quadrant. The two bandpass filter outputs of the first stage appear in the UR and LL quadrants, while the decimated lowpass output falls at the LR corner of the UL quadrant. The three output images of the second stage follow the same format in the UL quadrant.

This paper is organized as follows. In Section 2 we discuss the mathematical requirements imposed by invertibility and describe

a frequency-domain method for designing the required transfer functions. In Section 3 we illustrate the process with an example. Finally we discuss the technique and its application.

2. FILTER DESIGN

2.1 Invertibility Requirements

The constraints the filters must satisfy [1], illustrated for the $k = 2$ case, are (1) flat system response,

$$|H_0(u, v)|^2 + |L_0(u, v)|^2 \left[\begin{array}{l} |L_1(u, v)|^2 + |B_1(u, v)|^2 \\ + |B_2(u, v)|^2 \end{array} \right] = 1 \quad (1)$$

(2) recursion

$$\left| L_1\left(\frac{u}{2}, \frac{v}{2}\right) \right|^2 = \left| L_1\left(\frac{u}{2}, \frac{v}{2}\right) \right|^2 \left[\begin{array}{l} |L_1(u, v)|^2 + |B_1(u, v)|^2 \\ + |B_2(u, v)|^2 \end{array} \right] \quad (2)$$

and (3) no aliasing upon subsampling,

$$L_1(u, v) = 0 \quad \text{for } s > f_N/2 \quad (3)$$

where $f_N = 1/2\Delta x$ is the folding frequency, and $s = \sqrt{u^2 + v^2}$ is the radial variable in frequency space.

A sufficient condition for Eq. (2) to hold is that the decomposition/reconstruction filter bank has unity gain,

$$|L_1(u, v)|^2 + |B_1(u, v)|^2 + |B_2(u, v)|^2 = 1 \quad (4)$$

From this Eq. (1) implies that the pre- and post-processing steps must also have unity gain, that is,

$$|H_0(u, v)|^2 + |L_0(u, v)|^2 = 1 \quad (5)$$

This defines a special case for which it is simple to construct suitable transfer functions.

2.2 The Lowpass Filters

We define a general, smooth-edged, one-dimensional lowpass transfer function using the raised cosine,

$$LP(a, b, f) = \begin{cases} 1 & ; \quad f \leq a \\ \sqrt{\frac{1}{2} \left[1 + \cos \left[\pi \left(\frac{f-a}{b-a} \right) \right] \right]} & ; \quad a < f < b \\ 0 & ; \quad f \geq b \end{cases} \quad (6)$$

where a and b are parameters that establish the band limits. The transfer function of the two-dimensional lowpass filter used in the preprocessing stage is given by

$$L_0(u, v) = LP(f_2, f_N, s) \quad (7)$$

the lowpass filter used in the iterated stages is

$$L_1(u, v) = LP(f_1, f_N/2, s) \quad (8)$$

and f_1 and f_2 are design parameters. These transfer functions are shown in Figures 3 and 4.

2.3 The Highpass Filter

We define a smooth-edged one-dimensional highpass transfer function, again using the raised cosine,

$$HP(a, b, f) = \begin{cases} 0 & ; \quad f \leq a \\ \sqrt{\frac{1}{2} \left[1 - \cos \left[\pi \left(\frac{f-a}{b-a} \right) \right] \right]} & ; \quad a < f < b \\ 1 & ; \quad f \geq b \end{cases} \quad (9)$$

Then the two-dimensional highpass filter transfer function is given by

$$H_0(u, v) = HP(f_2, f_N, s) \quad (10)$$

This transfer function is shown in Figure 3.

2.4 The Steerable Bandpass Filters

The k directional bandpass filters used in the iterated stages are given by

$$B_m(u, v) = HP(f_1, f_N/2, s) \cos^{k-1}(\theta - m\pi/k) \quad (11)$$

where $m = 0, \dots, k-1$, and

$$\theta = \tan^{-1}(v/u) \quad (12)$$

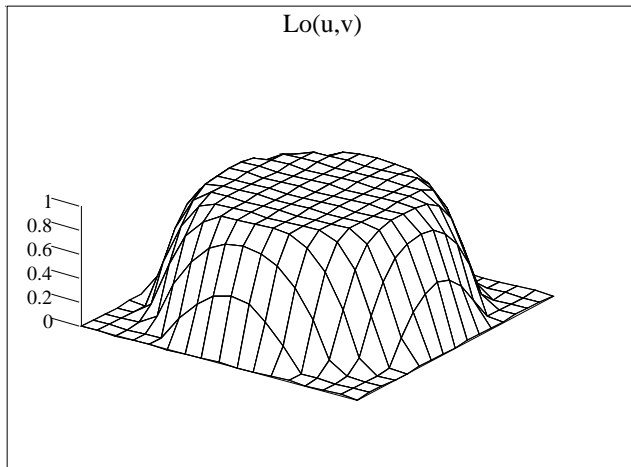
is the angular variable in frequency space.

3. DESIGN EXAMPLE: $k = 2$

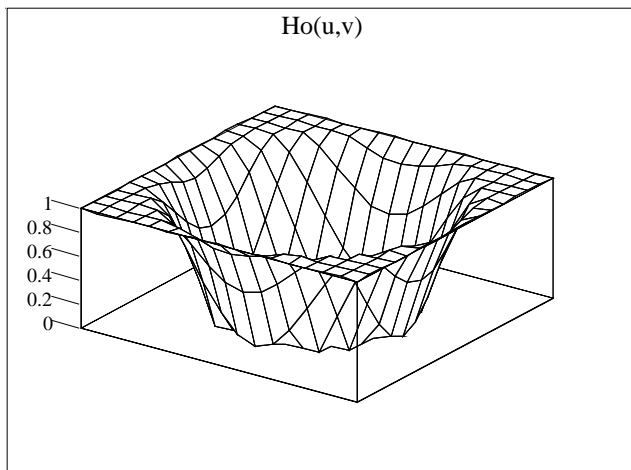
The $k=2$ case, illustrated in Figure 1, requires two bandpass filters, and these approximate first derivatives in the x - and y -directions. Their transfer functions are shown in Figure 4. In this example, the parameters f_1 and f_2 were chosen as 0 and $5/8 f_N$, respectively.

The sum of squared magnitudes of the two bandpass filters forms a circularly symmetric highpass filter. The lowpass filter, $L_1(u, v)$, passes the low frequency components that fall in the central core of that filter. Finally, the highpass filter, $H_0(u, v)$, passes the high frequency information that falls in the "corners" of the (rectangular) frequency plane. Thus all energy, regardless of frequency, is passed to the output with unity gain.

The output of the $L_1(-u,-v)$ decomposition filter is downsampled by a factor of two, but this is aliasing-free since all energy of frequency high enough to be aliased has been removed by $L_1(-u,-v)$. Upon reconstruction, the down- and up-sampled signal in branch 4 (in which every second sample is zero) is interpolated by $L_1(u,v)$.

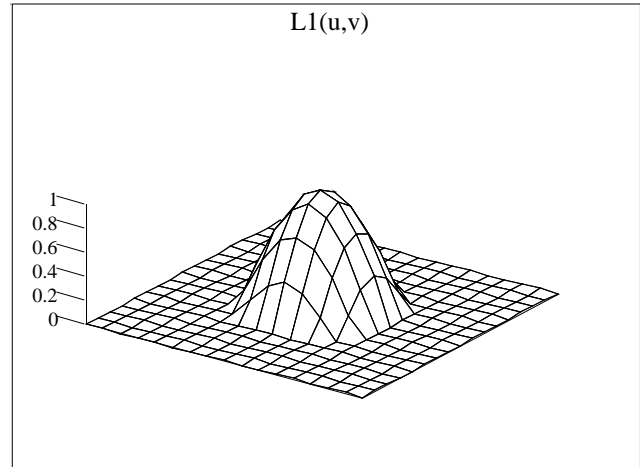


L0

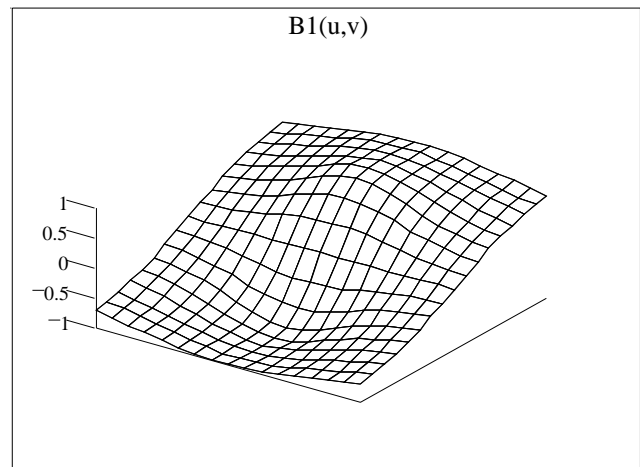


H0

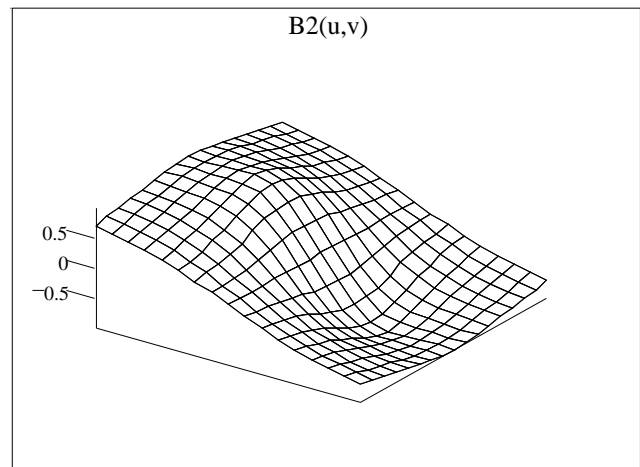
Figure 3. The transfer functions of the pre- and post-processing filters used in Figure 1.



L1



B1



B2

Figure 4. The transfer functions of the iterated filters used in Figure 1.

4. IMPLEMENTATION

The filter bank decomposition can be implemented in either the spatial or frequency domain. Normally the decomposition filters would be designed in the frequency domain, since the conceptual process is one of segmenting that domain with a set of transfer functions that sum to unity. Convolution kernels can be obtained by inverse Fourier transformation, provided that round-off and kernel truncation errors are kept under control [6].

With a large number of filters in the bank (i.e., large k), one might do well to implement the design process (i.e., determination of f_1 and f_2) in the frequency domain. A test image could be transformed, and the N component images inverse transformed initially to guide the design process. The output image would then have to be inverse transformed each time a new design was specified. The final design could be implemented with convolutions.

5. SUMMARY

The filter design technique presented here can produce implementations for steerable pyramid transforms of any degree. These can be used, for example, in image enhancement applications [7]. They are accurately invertible (perfect reconstruction filter banks), and the required filtering can be implemented in either the spatial or frequency domain. For spatial domain implementations the convolution kernels can be truncated to reasonable size without introducing serious approximation error.

Images formed by optical means are subject to the OTF of the image-forming lens. For microscopes this is, at best, a circular transfer function that is band-limited at $f_c = 2NA/\lambda$. Thus, if the image is not undersampled (i.e., if $f_N = 1/2\Delta x > f_c$), there will be no information outside the circle of radius f_N in the frequency plane, and the pre- and post-processing steps can perhaps be eliminated.

Implemented with the WiT software package [8,9], the transform executes with reasonable speed on a Pentium processor. This provides a convenient testbed for the development of image processing applications using the steerable pyramid transform.

6. ACKNOWLEDGEMENTS

This research was supported in part by the National Cancer Institute under Grant No. 1R43 CA76896-01, and by the National Institute of Child Health and Human Development under Grant No. 1R43 HD33658-01.

7. REFERENCES

- [1] E. P. Simoncelli, and W. T. Freeman, "The Steerable Pyramid: A Flexible Architecture for Multi-scale Derivative Computation," *Proc. ICIP-95*, 444-447.
- [2] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable Multiscale Transforms," *IEEE Trans. IT-38*(2):587-607, 1992.
- [3] P. J. Burt, and E. H. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Trans. C-31*:532-540, 1983.
- [4] R.H. Bamberger and M. J. T. Smith, "A Filter Bank for the Directional Decomposition of Images: Theory and Design," *IEEE Trans. SP-40*(4):882-893, 1992.
- [5] S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. PAMI-11*:674-693, 1989.
- [6] K. R. Castleman, *Digital Image Processing*, Prentice-Hall, 1996.
- [7] Q. Wu, M. A. Schulze, and K. R. Castleman, "Steerable Pyramid Filters for Selective Image Enhancement Applications," To appear in *Proceedings of ISCAS '98*, 1998.
- [8] Logical Vision, Ltd., 4299 Canada Way, Suite 265, Burnaby, B.C. Canada V5G 1H3,
- [9] Logical Vision, Ltd., <<http://www.logicalvision.com>>