An Orthogonal Oriented Quadrature Hexagonal Image Pyramid

Andrew B. Watson
Albert J. Ahumada, Jr., Ames Research Center, Moffett Field, California

December 1987

NASA
National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

AN ORTHOGONAL ORIENTED QUADRATURE HEXAGONAL IMAGE PYRAMID

Andrew B. Watson and Albert J. Ahumada, Jr.

NASA Ames Research Center
Perception and Cognition Group

Abstract

We have developed an image pyramid with basis functions that are orthogonal, self-similar, and localized in space, spatial frequency, orientation, and phase. The pyramid operates on a hexagonal sample lattice. The set of seven basis functions consist of three even high-pass kernels, three odd high-pass kernels, and one low-pass kernel. The three even kernels are identical when rotated by 60° or 120°, and likewise for the odd. The seven basis functions occupy a point and a hexagon of six nearest neighbors on a hexagonal sample lattice. At the lowest level of the pyramid, the input lattice is the image sample lattice. At each higher level, the input lattice is provided by the low-pass coefficients computed at the previous level. At each level, the output is subsampled in such a way as to yield a new hexagonal lattice with a spacing $\sqrt{7}$ larger than the previous level, so that the number of coefficients is reduced by a factor of 7 at each level. We discuss the relationship between this image code and the processing architecture of the primate visual cortex.

Introduction

A digital image is usually represented by a set of two-dimensionally periodic spatial samples, or pixels. Many schemes exist to transform these pixels into alternative image codes that may be useful for compression or progressive transmission. subband codes are a class of transform in which the image is partitioned into sub-images corresponding to separate bands of resolution or spatial frequency (Vetterli, 1984; Woods and O'Neil, 1986). Closely related are pyramid codes, in which each band-pass sub-image is sub-sampled by a common factor, so that the number of pixels in each level of the pyramid is reduced by that factor relative to the preceding level (Tanimoto and Pavlidis, 1975; Burt and Adelson, 1983, Watson, 1986). Several schemes have been devised that also partition the image by orientation. These include quadrature mirror filters (Vetterli, 1984; Woods and O'Niel, 1986; Gharavi and Tabatabai, 1986; Mallat, 1987), and a pyramid modeled on human vision (Watson, 1987a,b). Recently, a number of orthogonal pyramid codes have been developed (E. H. Adelson, Eero Simoncelli, and Rajesh Hingorani , Orthogonal pyramid transforms for image coding, *SPIE Proceedings on Visual Communication and Image Processing II*, 1988). These have the virtues that they are invertible, that they preserve the total number of coefficients, and that they allow simple forward and inverse transformation algorithms.

We are interested in image codes that share properties with the coding scheme used by the primate visual cortex (A. B. Watson, Cortical algotecture, in *Vision: Coding and Efficiency*, C. B. Blakemore, Ed., Cambridge University Press, Cambridge England, 1988). These properties include a subband structure, relatively narrow-band tuning in both spatial frequency and orientation, relatively high spatial localization, both odd and even (quadrature) kernels, and self-similarity. We have also been intrigued by the fact that the image sample lattice in primate vision is approximately hexagonal, rather than rectangular. Guided by these observations, we have derived an orthogonal oriented quadrature hexagonal image pyramid.

Our code is a shift-invariant linear transformation, in which each new coefficient is a linear combination of image samples. The linear combination can be defined by a kernel of weights specifying the spatial topography of the linear combination. We have considered kernels that occupy a point and the hexagon of six nearest neighbors on a hexagonal lattice.

Constraints

We have derived a set of kernels under the following constraints:

(1) The kernels are expressed on a hexagonal sample lattice.
(2) There are seven mutually orthogonal kernels, one low-pass and six high-pass.
(3) Each kernel has seven weights (taps) corresponding to a point and its six nearest neighbors in the hexagonal lattice.
(4) The low-pass kernel has equal values at all taps.
(5) Two high-pass kernels have an axis of symmetry running through the center sample and between samples on the outer ring (at an angle of 30°).
(6) Of these two kernels, one is even about the axis of symmetry, the other is odd.
(7) The remaining four high-pass kernels are obtained by rotating the odd and even kernels by 60° and 120°.
(8) Each kernel has a norm (square root of sum of squares of taps) of one.

With respect to constraint (5), we have determined that there is no solution when the common axis of symmetry is at 0° (on the sample lattice of the outer ring). Note also that constraints (2) and (4) oblige the even kernels, as well as the odd, to have zero DC response (the weights sum to 0).

Under the symmetry constraints, the kernel coefficients can be written as shown in Fig. 1.

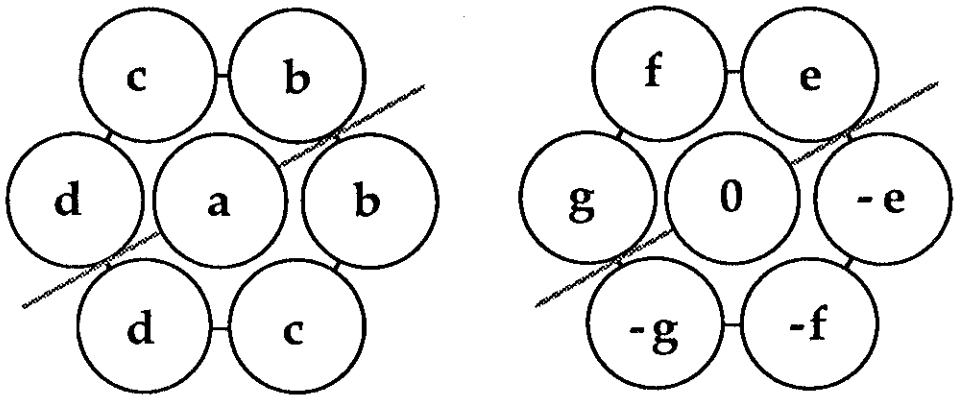

Fig. 1. Even and odd high-pass kernels with symmetry axis at 30°.

One even and one odd kernel are shown. The low-pass kernel (not shown) is simply a constant *h*at each tap. We construct a set of seven equations in these seven unknowns that express the constraints of orthogonality and unit norm. They are:

$$a^2 + 2b^2 + 2c^2 + 2d^2 = 1 \qquad \text{(unit norm)} \qquad [1]$$

$$2e^2 + 2f^2 + 2g^2 = 1 \qquad \text{(unit norm)} \qquad [2]$$

$$a + 2b + 2c + 2d = 0 \qquad (\perp \text{ to low-pass}) \qquad [3]$$

$$a^2 + b^2 + d^2 + 2bc + 2cd = 0 \qquad (\perp \text{ to self-rotation}) \qquad [4]$$

$$a^2 + 2bc + 2bd + 2cd = 0 \qquad (\perp \text{ to self-rotation}) \qquad [5]$$

$$e^2 + g^2 - 2ef - 2fg = 0 \qquad (\perp \text{ to self-rotation}) \qquad [6]$$

$$2eg - 2ef - 2fg = 0 \qquad (\perp \text{ to self-rotation}) \qquad [7]$$

Subtracting equations [4] and [5], and [6] and [7], shows that

$$b = d \qquad [8]$$

$$e = g \qquad [9]$$

Thus while not explicitly assumed, we see that both odd and even filters must also

be symmetrical about the 120° axis.

Further simplifications lead to the following solution for the coefficients of the odd filter:

$$e = \sqrt{2}/3 \qquad [10]$$

$$f = e/2 = \frac{1}{\sqrt{2}\,3} \qquad [11]$$

For the even filter, we find:

$$a = \sqrt{2/7} \qquad [12]$$

But two solutions emerge for b and c:

$$b = \frac{-(1 + 1/\sqrt{7})}{\sqrt{2}\,3} \qquad [13]$$

$$c = \frac{(2 - 1/\sqrt{7})}{\sqrt{2}\,3} \qquad [14]$$

and

$$b = \frac{(1 - 1/\sqrt{7})}{\sqrt{2}\,3} \qquad [15]$$

$$c = \frac{-(2 + 1/\sqrt{7})}{\sqrt{2}\,3} \qquad [16]$$

We will call the first solution the even filter of type 0, and the second solution, type 1. The three kernels are shown in Fig. 2.

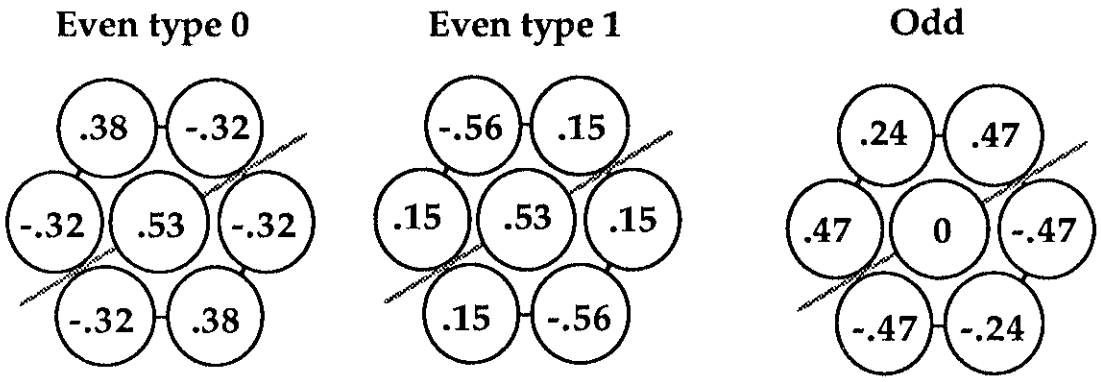

Fig. 2. Values for the two types of even kernel and one odd kernel.

The value of each coefficient h in the low-pass kernel is given directly by the unit norm constraint,

$$h = 1/\sqrt{7} \qquad [17]$$

Filter spectra

One of our objectives was to create subband filters that were somewhat narrowband and oriented. The filter spectra are easily derived. Each kernel consists of a central impulse at the origin, surrounded by 3 pairs of symmetric impulses. These transform in the frequency domain into a constant plus three sinusoids at angles of 0°, 60°, and 120°. The constant is the value of the central coefficient, while each sinusoid has an amplitude twice that of the corresponding coefficient. For the even kernels, the sinusoids are in cosine phase, for the odd kernels, they are in sine phase. The example spectra shown in Fig. 3 demonstrate from their half amplitude response that they are oriented and high-pass. In the pyramid they will become band-pass through convolution with the low-pass kernel at preceding levels.

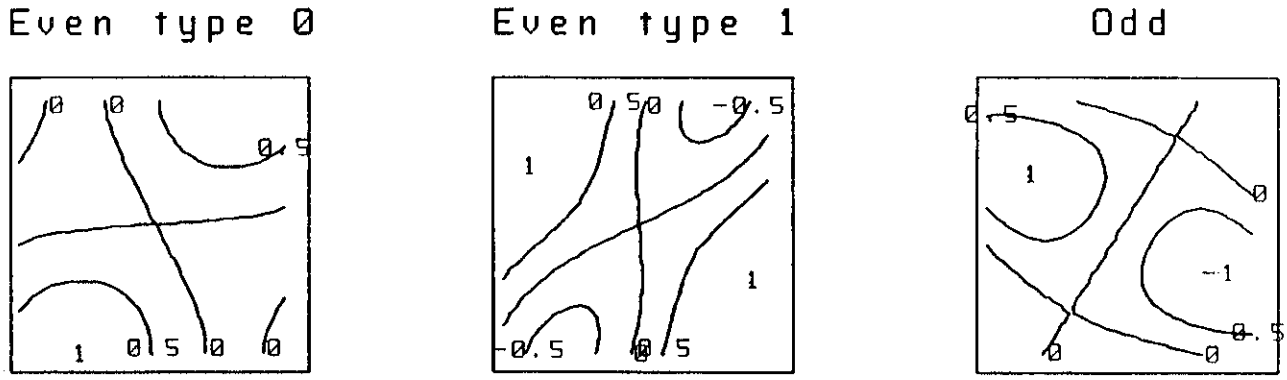

Fig. 3. Spectra of the two types of even kernel and the odd kernel. The origin is at the center of each figure and the spectrum extends to plus and minus 1. These are contour plots of continuous spectra. The discrete spectrum would have a hexagonal shape and a hexagonal sample lattice.

Axes of symmetry and orientation

We define the *orientation* of a kernel as the orientation of the peak of the frequency spectrum, that is, the orientation of a sinusoidal input at which the kernel gives the largest response. An interesting feature of the resulting kernels is that while the axis of symmetry was fixed at 30°, the orientation of the type 0 even kernel is actually orthogonal to this axis at 120°. This places its orientation axis on the hexagonal lattice. In contrast, the orientation of the type 1 even kernel and the odd kernel are equal to the initial axis of symmetry at 30°. Thus if it is desired to have quadrature pairs with equal orientation, the type 1 even kernel must be used.

Subsampling

One virtue of the scheme we have described is that it leads directly to an oriented resolution pyramid, as illustrated in Fig. 4.

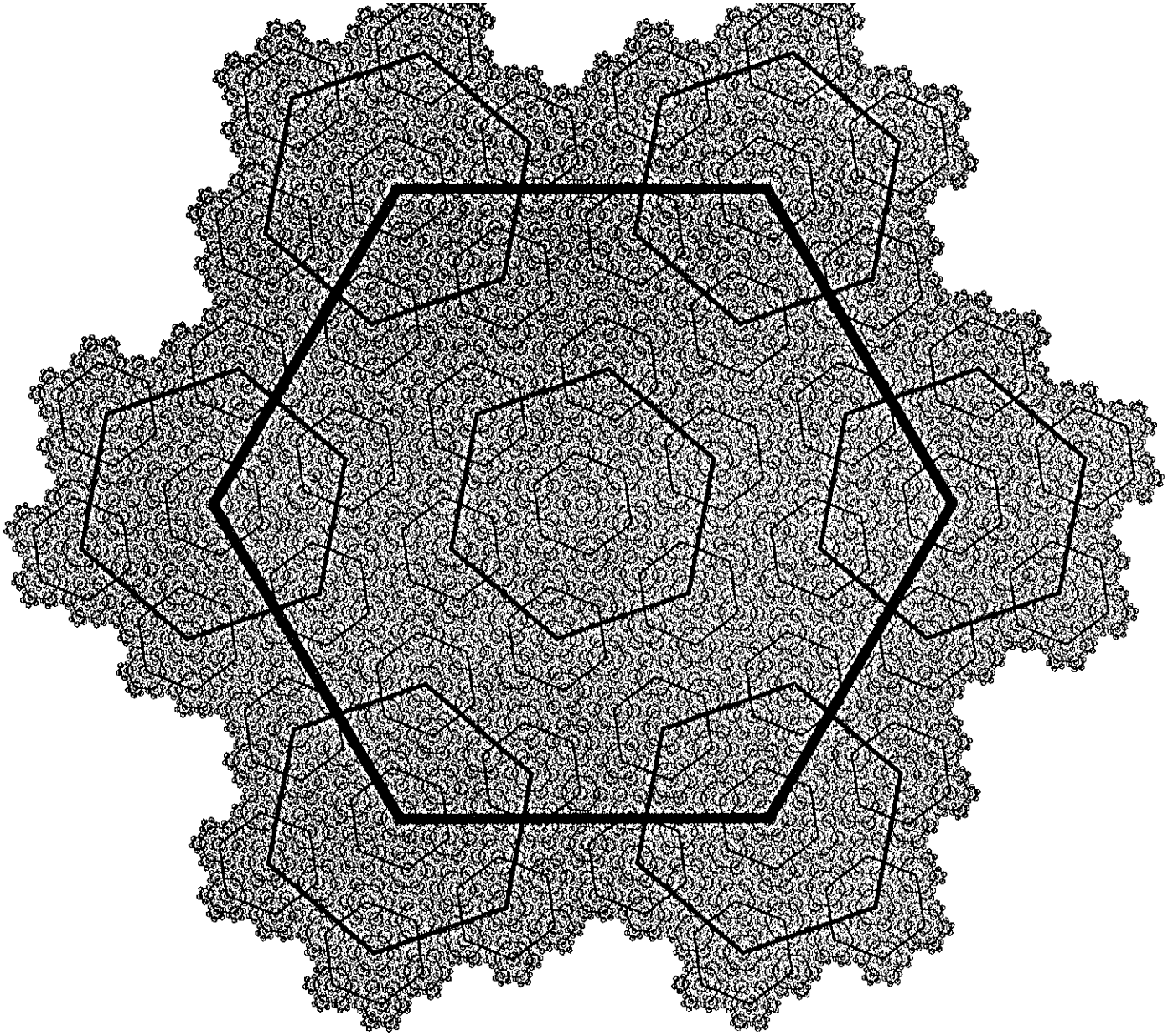

Fig. 4. Construction of the hexagonal pyramid. The image sample lattice is given by the vertices of the smallest hexagons. At each level, sub-images are generated by application of the kernels to the low-pass coefficients from the previous level.

This hexagonal fractal was constructed by first creating the largest hexagon, then placing at each of its vertices a hexagon rotated by $\tan^{-1}(\sqrt{3}/5) \approx 19.1°$ and scaled by $1/\sqrt{7}$. The same procedure is then applied to each of the smaller hexagons, down to some terminating level. The image sample lattice is then a finite-extent periodic sequence with a hexagonal sample lattice defined by the vertices of the smallest hexagons. The sample lattice has 7^6 points, the same as a rectangular lattice of 343^2. The perimeter of this "Gosper flake" is a "Koch curve" with a fractal dimension of $\log 3 / \log \sqrt{7} \approx 1.19$ (Mandelbrot, 1983, p. 46). The program used to create this image is given in Appendix 1.

The hexagonal image sample lattice is tesselated with hexagons with unit radius. Each of the 7 kernels is applied in each hexagon, yielding seven new sub-images, six highpass and one lowpass, each with one seventh as many samples as the original. The six high-pass sub-images form level 0 of the pyramid. The next level is created by again tesselating the plane with hexagons of radius $\sqrt{7}$ whose vertices correspond to the centers of the hexagons at the lower level. The seven kernels are applied to the low-pass coefficients derived at the earlier level. This yields seven new sub-images, each a factor of seven smaller than the sub-images at level 0. This process is repeated until a level is reached at which each sub-image has one sample.

While an image shape like that in Fig. 4 is very natural for this code, any shape that is one period of a hexagonally periodic sequence can be exactly encoded if the number of samples is equal to a power of seven. This includes, for example, a parallelogram with sides of length a power of seven samples. Below we show how the code may be applied to a conventional rectangular image.

The sub-sampling at each level can be formalized as follows (Dudgeon and Mersereau, 1984). The original hexagonal sampling lattice can be represented by a sampling matrix **H**,

$$\mathbf{H} = \begin{bmatrix} 1 & 1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix} \qquad [18]$$

The column vectors of this matrix map from sample to sample, and the location of any sample can be expressed as $\mathbf{x} = (x,y)$,

$$\mathbf{x} = \mathbf{H}\,\mathbf{r} \qquad [19]$$

where **r** is an integer vector. Let $\mathbf{S}_n$ be the sampling matrix at level n. Since the sample at each level must be a subset of those at the previous level, the column vectors of $\mathbf{S}_{n+1}$ must be integer linear combinations of the column vectors of $\mathbf{S}_n$. Thus

$$\mathbf{S}_{n+1} = \mathbf{S}_n\mathbf{M} \qquad [20]$$

where $\mathbf{M}$ is an integer matrix. Further, the columns of $\mathbf{S}_{n+1}$ must be $\sqrt{7}$ longer than the columns of $\mathbf{S}_n$ (corresponding to the increasing radii of the hexagons at each successive level). And finally, because the determinant of a sampling matrix determines the factor by which the density of samples is reduced, we know that

$$\det(\mathbf{M}) = 7 \qquad [21]$$

Two matrices which satisfy these conditions are:

$$\mathbf{M}_0 = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} \qquad [22]$$

$$\mathbf{M}_1 = \begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix} \qquad [23]$$

These generate the only two possible sub-samplings from one level to the next. Then $\mathbf{S}_n$ can be constructed in various ways, the three most obvious being

$$\mathbf{S}_n = \mathbf{H}\,\mathbf{M}_0^n \qquad [24]$$

and

$$\mathbf{S}_n = \mathbf{H}\,\mathbf{M}_1^n \qquad [25]$$

and

$$\mathbf{S}_n = \mathbf{H}\,\mathbf{M}_0\mathbf{M}_1\mathbf{M}_0\mathbf{M}_1\ldots \quad (n \text{ terms}) \qquad [26]$$

The first scheme (used in Fig. 4) causes a rotation of $\tan^{-1}(\sqrt{3}/5) \approx 19.1°$ in the sample lattice at each level, as does the second scheme, while the third scheme alternates between rotations of 19.1° and -19.1°.

Skewed coordinates

It is well known that hexagonal samples on a cartesian plane can also be viewed as rectangular coordinates on a coordinate frame in which one axis is skewed by 60° (Fig. 6A) (Peterson and Middleton, 1962; Mersereau, 1979).

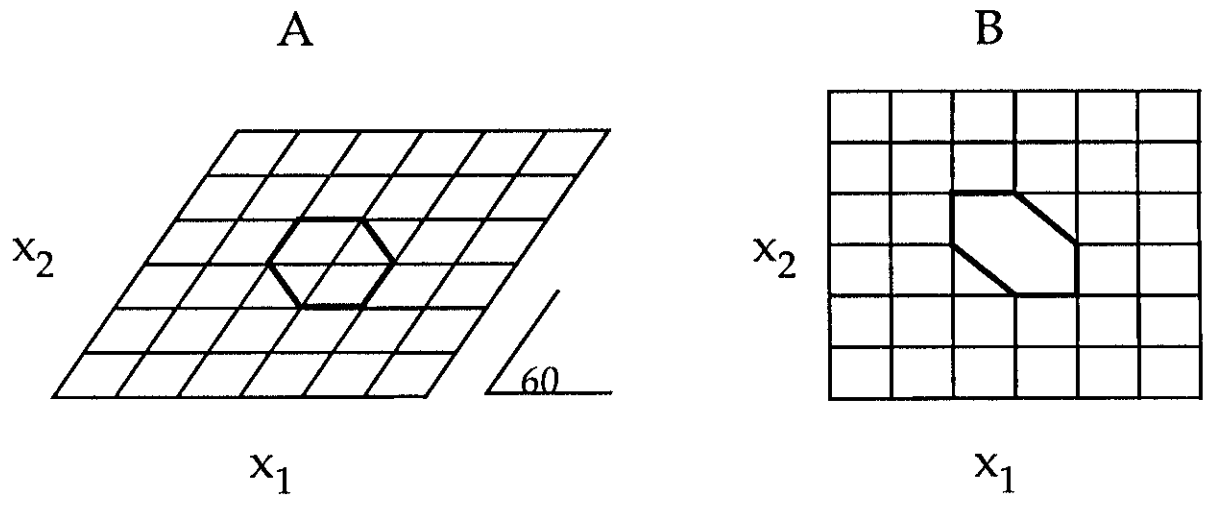

Fig. 6. A) Hexagonal lattice represented as skewed rectangular coordinates. B) De-skewed rectangular coordinates. The hexagon is distorted into an oblique lozenge.

In this coordinate scheme, the sampling matrices are even simpler. They are the same as above (Eq.s 24, 25, and 26) except that we drop the matrix $\mathbf{H}$ from each expression.

This leads to a natural method for application of this coding scheme to conventional rectangular images. When the skewed coordinates are "de-skewed" (Fig. 6B), the hexagon is distorted into an oblique lozenge. The orthogonal pyramid may then be constructed using these lozenges as the shape for each kernel. The kernels will no longer be rotationally symmetric, but for some purposes this may be unimportant. As before, exact coding will be possible so long as the sides of the rectangle are a power of seven.

Biological image coding

One likely role of the primate visual cortex is to encode the retinal image in components that are less correlated than the image pixels themselves. The scheme we have described provides a model for how this might be done. In this context, the initial samples (indicated by the vertices of the smallest hexagons in Fig. 4) correspond to the receptive field centers of retinal ganglion cell inputs. Each hexagon defines the receptive field of a single cortical unit. The coefficients of each basis function describe the weights with which each ganglion cell contributes to the response of the cortical cell. The basis functions defined on the smallest hexagons correspond to the cells tuned to the highest spatial frequencies. Each subsequent level of the pyramid corresponds to cells tuned to lower and lower frequencies. The low-pass basis functions at each level correspond to un-oriented pooling units, which in turn are used to create the high-pass units at the next level. These pooling units may correspond to actual cells, or may simply define which ganglion cells contribute inputs to the high-pass units at each level.

Elsewhere we have introduced the term *chexagon* (cortical hexagon) to describe the generic scheme of construction of cortical receptive fields through combination of retinal ganglion cell inputs laid out on a hexagonal lattice (A. B. Watson, Cortical

algotecture, in *Vision: Coding and Efficiency*, C. B. Blakemore, Ed., Cambridge University Press, Cambridge England, 1988). The present chexagon scheme agrees with what is known about cortical cells in several respects. The high-pass filters are tuned for both spatial frequency and orientation. The input lattice of ganglion cells is known to be approximately hexagonal, at least in the foveal region. The shape of the one-dimensional pass-band of each filter, when mutiplied by the pass-band of the ganglion cells, is similar to that of cortical cells. Finally, cortical cells are believed to form quadrature pairs, like the odd and even basis functions described here.

There are on the other hand a number of respects in which this scheme appears to differ from cortical coding. First, the frequency tuning functions of our filters are oriented in the sense of having a strongest response at one orientation, but they have a second lobe of response (of opposite sign) at the orthogonal orientation. Two-dimensional mapping of frequency tuning functions in cortical cells occasionally show such secondary lobes (De Valois, Yund, and Hepler, 1982), but they do not appear to be common. Second, the units we describe change in scale by $\sqrt{7}$ at each level, which might yield rather fewer different scales than are commonly supposed. Third, the 19.1° rotation of the axis of orientation at each scale reduces the degree of rotation invariance of the code, though rotational invariance is not known to hold for the cortical code. Fourth, the tuning functions produced by our scheme are broader in orientation than in spatial frequency. While subject to some debate, it is believed that this is opposite to the aspect ratio of cortical cells.

Finally, the precise crystaline structure of this code is clearly different from the biological heterogeneity of visual cortex. Nonetheless, the cortex is highly regular, and a scheme like ours may be the canonical form from which the actual cortex is a developmental perturbation. These issues are discussed at greater length elsewhere (A. B. Watson, Cortical algotecture, in *Vision: Coding and Efficiency*, C. B. Blakemore, Ed., Cambridge University Press, Cambridge England, 1988). Perhaps the best summary is that while this scheme may not describe exactly the cortical encoding architecture, it is an example of the form such an architecture might take.

References

P. J. Burt and E. H. Adelson, The Laplacian Pyramid as a Compact Image Code, *IEEE Transaction on Communications*, COM-31, No. 4, April 1983, 532-540.

R. L. De Valois, E. W. Yund, and H. Hepler, The orientation and direction selectivity of cells in macaque visual cortex, *Vision Research*, 22, 1982, 531-544.

D. A. Dudgeon and R. M. Mersereau, *Multidimensional digital signal processing*, Englewood Cliffs, NJ, Prentice-Hall, 1984.

H. Gharavi, Ali Tabatabai, Sub-band Coding of Digital Images Using Two-Dimensional Quadrature Mirror Filtering, *SPIE Proceedings on Visual Communication and Image Processing*, **707**, 51-61, 1986.

S. G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, *GRASP Lab Technical Memo* **MS-CIS-87-22**, University of Pennsylvania Dept. of Computer Information Science, 1987.

B. B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, New York, 1983.

R. M. Mersereau, The processing of hexagonally sampled two-dimensional signals, *Proceedings of the IEEE*, 67, No. 6, 1979, 930-949.

D. P. Petersen and D. Middleton, Sampling and reconstruction of wave-number limited functions in N-dimensional Euclidean spaces, *Inform. Contr.*, 5, 1962, 279-323.

S. Tanimoto and T. Pavlidis, A hierarchical data structure for picture processing,

Computer Graphics and Image Processing, 4, 1975, 104-119.

M. Vetterli, Multi-Dimensional sub-band coding: some theory and algorithms, *Signal Processing*, **6**, 97-112, 1984.

A. B. Watson, Ideal shrinking and expansion of discrete sequences, *NASA Technical Memorandum 88202*, January 1986.

A. B. Watson, The cortex transform: Rapid computation of simulated neural images, *Computer Vision, Graphics, and Image Processing*, **39**, 311-327, 1987a.

A. B. Watson, Efficiency of an image code based on human vision, *Journal of the Optical Society of America A*, December 1987b.

J. W. Woods and S. D. O'Neil, Subband coding of images, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-34, No. 5, October 1986, 1278-1288.

Appendix 1

The following is a program in the Postscript language to draw the chexagon pyramid in Fig. 4. The number of levels drawn is determined by the variable *maxdepth*. On an Apple laser printer, a *maxdepth* of 3 takes about 2 minutes to print. Each greater depth will take a factor of 7 longer.

```
statusdict /jobname (Beau fracthex.ps) put
/#copies 1 def
/timezero usertime def
/showSTATUS { = usertime timezero sub 1000 idiv = (Secs) = flush} def

/depth 0 def
/maxdepth 3 def                                   % maximum levels
/latticeRot 3 sqrt 5 atan def                     % lattice rotation angle
/root7 1 7 sqrt div def                           % scale change between levels
/negrot {/latticeRot latticeRot neg def} def
/down {/depth depth 1 add def } def               % increments depth
/up {/depth depth 1 sub def } def                 % decrements depth
/inch {72 mul} def                                % scale to inches

/hexside {60 rotate 1 0 lineto currentpoint translate } def   % draw one side of a hexagon

/drawhex                                          % draw unit hexagon
{ gsave
-60 rotate 1 0 moveto 60 rotate currentpoint translate        % move to first vertex
5 { hexside } repeat                              % draw 5 sides
closepath stroke                                  % draw sixth side
grestore } def

/vertex    % angle is on stack        % go to vertex at angle, draw hexagon pyramid
{/angle exch def
gsave
angle rotate 1 0 translate angle neg rotate
fracthex
grestore
} def
```

```
/fracthex                                   % draw hexagon pyramid
{gsave
root7 dup scale                             % reduce scale by root 7
2 72 div setlinewidth
down negrot latticeRot rotate drawhex       % move down one level, rotate lattice, draw hex
depth maxdepth le                           % test if at max level
    {fracthex                               % recursive call to fracthex
    0 60 300 { vertex } for                 % call vertex at each vertex
    } if
up negrot grestore
} def

gsave                                       % main program
4.25 inch 5.5 inch moveto currentpoint translate   % set origin
6 inch 6 inch scale                         % set global scale
latticeRot neg rotate                       % set initial orientation
1 setlinejoin
fracthex                                    % do it
grestore
1 inch 1 inch moveto
/Palatino-Roman findfont 34 scalefont setfont   % label
(Chexagon Pyramid) show
showpage
(Elapsed time) showSTATUS
```

National Aeronautics and Space Administration

Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM-100054		

4. Title and Subtitle	5. Report Date
An Orthogonal Oriented Quadrature Hexagonal Image Pyramid	December 1987
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Andrew B. Watson and Albert J. Ahumada, Jr.	A-88049
	10. Work Unit No. 506-47

9. Performing Organization Name and Address	11. Contract or Grant No.
Ames Research Center Moffett Field, CA 94035	
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address	Technical Memorandum
National Aeronautics and Space Administration Washington, DC 20546-0001	14. Sponsoring Agency Code

15. Supplementary Notes

Point of Contact: Andrew B. Watson, Ames Research Center, MS 239-3, Moffett Field, CA 94035, (415) 694-5419 or FTS 464-5419

16. Abstract

We have developed an image pyramid with basis functions that are orthogonal, self-similar, and localized in space, spatial frequency, orientation, and phase. The pyramid operates on a hexagonal sample lattice. The set of seven basis functions consist of three even high-pass kernels, three odd high-pass kernels, and one low-pass kernel. The three even kernels are identical when rotated by 60° or 120°, and likewise for the odd. The seven basis functions occupy a point and a hexagon of six nearest neighbors on a hexagonal sample lattice. At the lowest level of the pyramid, the input lattice is the image sample lattice. At each higher level, the input lattice is provided by the low-pass coefficients computed at the previous level. At each level, the output is subsampled in such a way as to yield a new hexagonal lattice with a spacing $\sqrt{7}$ larger than the previous level, so that the number of coefficients is reduced by a factor of 7 at each level. We discuss the relationship between this image code and the processing architecture of the primate visual cortex.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Image processing Image coding Sub-band Hexagonal sampling Human vision	Unclassified - Unlimited Subject Category - 59

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	17	A02

NASA FORM 1626 OCT 86