

A Model of Form and Texture Processing in V4 Based on Recursive Filtering

D. G. Lane
December 7, 1995

Visual responses in area V4 span a diverse set of stimuli. Originally considered a color area, later studies have reported responses to bars of varying widths and lengths, sinusoidal gratings, non-cartesian stimuli, textures and texture gradients, and more complex patterns. The diversity of receptive field properties defies attempts to formulate a single model of form and texture processing within V4. However, a unified model is necessary to ultimately understand how neurons in V4 interact to create a complete representation of complex scenes.

A biologically plausible model is proposed to explain some of the form and texture-processing characteristics of cells in V4. This model utilizes a bank of spatial frequency and orientation filters in V4, analogous to models of V1, that acts on the output of the V1 filter bank. The results of this filter-refilter operation are receptive field properties that resemble many of the properties observed in V4.

DESCRIPTION OF THE MODEL

The first issue to address by way of describing the recursive filters model is the effectiveness of the model. A filter-refilter operation sounds suspiciously redundant—at first glance, the refiltering operation would not seem to reveal any more information than the original filtering operation. This would be the case if the filter banks were only spatial frequency filters, or if only orientation filters were used. The key to the effectiveness of a filter-refilter operation is the interplay between spatial frequency filtering and orientation filtering.

For example, the two herringbone patterns in figure 1 are clearly distinct. However, the distribution of energy into oriented components is identical for both patterns: both distribute half of their energy in a 45 degree orientation, and the other half in a 135 degree orientation. The distribution of energy into spatial frequency bands is also identical between the two patterns. Thus, a bank of spatial frequency and orientation filters, such as V1, would be unable to discriminate these two patterns.

Consider the herringbone analysis system schematized in figure 2. One of the herringbone patterns is initially orientation filtered into four oriented maps. These four maps are then low-pass filtered, and then orientation refiltered. The result is a system of sixteen recursive oriented maps, and this particular herringbone pattern has energy distributed into two of these sixteen maps. The other herringbone pattern would have energy distributed into two other maps.

This basic configuration is expanded in figure 3 to depict the full recursive filters model. An initial bank of spatial frequency and orientation filters forms the first stage of processing, corresponding to visual processing through V1. The power from each of the resulting maps (the output of V1 complex cells) is then fed into the second stage. The second stage, representing V4, refilters the power maps with another bank of spatial frequency and orientation filters.

The recursive filter bank covers the same range of orientations as the initial filter bank. Only a restricted subset of the initial spatial frequencies is implemented in the recursive bank, though. Refiltering at a spatial frequency higher than the initial spatial frequency is not necessary. Simulations with several natural scenes indicate that spatial frequency energy in the recursive filter bank rarely distributes below two spatial frequency octaves of the initial spatial frequency, so the recursive filter bank's spatial frequency space is restricted to two octaves below the initial spatial frequency. Maps which are not implemented are greyed in figure 3.

The initial filter bank has a representation space that is four dimensional: two spatial dimensions, a spatial frequency dimension, and an orientation dimension. The recursive filter bank adds two additional dimensions: the recursive

spatial frequency and the recursive orientation. This expansion of the representation is partially offset by the decreased representation of the recursive spatial frequency space described above. The increased dimensionality of the recursive representation space could have a bearing on the topographic map of this space, as discussed below.

EVIDENCE FOR THE MODEL

One of the main assumptions of the recursive filters model is that the more complex recursive filter bank is located at a higher level in the visual processing stream than the initial filter bank. This is vital to the model, as the recursive filter bank can only account for complex receptive fields by acting on the output of the initial filter bank. The recursive filter bank is no more intrinsically complex than the initial filter bank; it will be argued below that both use similar local circuits to perform their respective computations. The additional complexity of the output of the recursive filter bank is due to the hierarchical arrangement of the system: the output of the initial filter bank is the input to the recursive filter bank.

The hierarchical arrangement of the visual system has been well-documented (Felleman and Van Essen 1983). Two major processing streams are identifiable: the parietal stream, which is thought to be involved in spatial localization and motion processing, and the occipito-temporal stream, which is thought to be involved in pattern and object recognition. The recursive filters model is a model of processing within the occipitotemporal stream. The hierarchy of the occipitotemporal stream begins in V1 and continues into V2, V4, and then into the inferotemporal complex.

If the processing in V4 is due to a refiltering of the information relayed from V1, inactivation of V1 should abolish activity in V4. This is important because other pathways have been identified from subcortical areas to V4. If V4 continues to operate without receiving information from V1, the recursive filters model is invalid. Girard et. al. (1992) have demonstrated that the inactivation of V1 with cryodes does abolish the visual responses of V4 cells, indicating that the output of V1 is necessary for V4 to respond.

Within the occipitotemporal stream, distinct processing streams have been identified based on anatomical evidence (DeYoe et. al. 1994). In V2, cytochrome oxidase stains for thick and thin stripes. The thick stripes are part of the parietal stream, but both the thin stripes and the interstripe regions are part of the occipitotemporal stream. This subdivision continues into V4, where distinct subregions can be discerned based on projections from either V2 thin stripes or interstripes. It is thought that the thin stripes and the V4 regions to which they project are involved with color processing, while the interstripes and the V4 regions to which they project are involved in form and texture processing. The recursive filters model applies specifically to the V2 interstripe to V4 subdivision. The model does not yet assign a specific role for V2. There is a bypass pathway directly from V1 to V4, but this consists mostly of the foveal representation of the visual field.

The range of the recursive spatial frequency maps is restricted relative to the range of the initial spatial frequency, as mentioned above. This aspect of the model makes a prediction concerning the relative sizes of receptive fields in V1 and V4. Since V4 is assumed to band-pass filter information from V1, the receptive field sizes of V4 would need to be larger than those in V1. Because the recursive filters in V4 band-pass to two octaves below the spatial frequencies in V1, the V4 receptive field size would need to be at least four times as large linearly as those in V1. Deemone and Sohein (1985) have observed that V4 receptive fields tend to be four to seven times as large as those in V1.

One of the motivations for developing the recursive filters model was a biologically plausible explanation for non-cartesian selective cells in V4. Some cells in V4 respond preferentially to concentric, radial, spiral, and hyperbolic stimuli (Gallant et. al. 1993). The recursive filters model in its present form provides a potential explanation for the selectivity from concentric, radial, and spiral stimuli, and extensions to the model may also account for hyperbolic selectivity.

The non-cartesian function of the recursive filters model can be understood by looking out why a single bank of spatial frequency and orientation filters can not discriminate radial and concentric stimuli. In figure 3, the initial filter bank demonstrates the decomposition of these stimuli into oriented components. Both of these stimuli have energy distributed equally in all available orientations. Thus, the initial filter bank is incapable of discriminating between the two stimuli.

The shape of the energy distributions within each initial oriented map does distinguish the two stimuli. The concentric stimulus has distribution consisting of a series of small parallel segments, arranged along a line that is orthogonal to the orientation of the smaller segments. The parallel stimulus has a single segment in each of the initial oriented maps.

When these distributions reach the recursive filter bank, this difference manifests itself in a distinct distribution among the recursive maps. The key is the distributions in the recursive oriented maps at the spatial frequency immediately lower than the original spatial frequency. The concentric stimulus has energy in these maps at an orientation that is orthogonal to each of the initial orientations. The radial stimulus has energy in these maps at an orientation parallel to the initial orientation. A linear summation of the appropriate recursive oriented maps results in a cell whose receptive field properties are selective for either radial or concentric stimuli.

In an analogous fashion, cells selective for spiral stimuli can be explained as a summation over the recursive oriented maps whose recursive orientation is at an oblique angle to the initial orientation. The sign of the oblique angle determines the polarity of spiral stimulus to which the cell would be most responsive. Most of the spiral sensitive cells observed by Gallant were not selective for spiral polarity, indicating that the model of these cells would require a summation over both oblique angle signs.

The performance of this model on non-cartesian stimuli was verified in a simulation experiment. An image processing library (APUtils, by Dr. Charles Anderson) was used to simulate both the initial and recursive filter banks. An input image was prepared consisting of either a single concentric stimulus or a single radial stimulus. This input image was then processed by the image processing software to create a set of twelve images representing the output from a bank of spatial frequency and orientation filters (3 spatial frequency levels and 4 orientation angles).

Each of these twelve images was then processed to create a power map, which is necessary to combine the components of both of the quadrature phases of the filtered images. The four power maps for the spatial frequency containing most of the stimulus' energy (the middle spatial frequency) were then used as inputs to the same routine that acted as the initial spatial frequency filter bank. The output from this step was then compared between the concentric and radial stimuli.

The distribution of energy in the recursive filter bank outputs was distinct between the two stimuli. The concentric stimulus had energy mostly in the "recursive orthogonal" maps; that is, the maps formed by recursively filtering at an orientation orthogonal to the initial orientation. The radial stimulus had energy mostly in the "recursive parallel" maps. This simulation experiment demonstrates

that the energy distributions among the recursive filter bank outputs are capable of classifying complex patterns.

Desimone and colleagues investigated receptive field properties of V4 cells using stimuli simpler than the non-cartesian stimuli (Desimone and Schein 1985, Desimone et. al. 1987). Using bars of that varied in length as well as width and orientation, they discovered many cells in V4 tuned to particular combinations of bar length and width. The distribution of lengths and widths was not highly correlated with receptive field size, indicating a receptive field with many subfields. The recursive filters model predicts receptive fields with many subfields, so it is natural to ask whether selectivity for bar length and width is in line with the recursive filters model.

A detector for bar length could result from the distribution of energy in the recursive filter maps. In particular, the width of a bar would correspond to the energy distribution in the initial oriented maps, and the length of the bar would correspond to the energy distribution in the recursive orthogonal oriented maps. Such a system would have various cells responsive to combinations of bar length and width.

Psychophysical evidence exists indicating that the visual system is capable of discriminating stimuli based on the local arrangement of features within each stimulus. Such as discrimination can occur via pop-out, indicating that a parallel mechanism is at work. Heathcote and Mawry (1993) looked at pop-out of an oddball which differed from the distractors only in the local arrangement of features. The results of this experiment indicate that a parallel mechanism is present in the visual system which is capable of distinguishing patterns based on the local arrangement of features. These experiments did not address the question of where such a system would be located. The recursive filters model provides a plausible explanation of this result.

The question of the types of local circuits responsible for the filtering properties in V1 is still not answered. Several models have been proposed (Hubel and Wiesel 1962; Ferster 1987; Somers et. al. 1995), but none can be unequivocally placed above the others. The recursive filters model does not specify which model of orientation filtering is more appropriate. Rather, the claim is made that V4 possesses local circuits similar to those found in V1.

Several studies support this assertion. Labeling of intralaminar GABAergic connections in several visual areas was compared in Kritzer et. al. 1992. The distribution in V4 was more expansive than in V1, but both were nearly circular.

Using biocytin to label the intrinsic lattice, Yoshioka et. al. (1992) found that horizontal connections in V4 were patchy, as in V1. The V4 patches spread out over a larger region of the cortex than the V1 patches, as was confirmed in Amir et. al. (1993). The V4 patches also tended to aggregate into clusters. This could be the sign of a more complex topographical arrangement in V4, corresponding to a larger dimensional representation space in V4.

PREDICTIONS OF THE MODEL

The most immediate predictions of the recursive filters model is the extended class of stimulus patterns that span the recursive representation space. Such stimuli include the non-cartesian stimuli, as well as stimuli created by amplitude modulating Gabor functions with lower spatial frequency Gabor functions. These amplitude modulated stimuli can also be assembled to create herringbone-like patterns.

In order to investigate the V4 mapping of this representation space, a method is needed to systematically generate stimuli in this space. This is possible because the filters implemented in the APUtills package are invertible. Thus, white noise patterns can be generated and used as inputs to the inverse filters. The result should be patterns which have distinct energy distributions in the recursive filter bank. One problem that needs to be resolved in order to successively generate these stimuli is the quadrature of the filters: the phases of the white noise could be randomly varied, but it is not clear that this will span the representation space uniformly.

Such stimuli can also be used in pop-out psychophysics experiments. Choosing two stimuli for a pop-out experiment involves ensuring that the two stimuli have equal distributions of energy in the initial filter banks. If this is the case, then evidence for parallel search of the two stimuli would indicate that processes in the visual system are going beyond the initial filter bank's features.

A more sophisticated question is opened but not resolved by the model. If this enhanced, six-dimensional representation space is being mapped in V4, there may be some topographical arrangement that could be detected. This topographical arrangement would be more complex than the pinwheels and fracture zones reported in V1 using optical imaging techniques. The V1 topography results from a mapping of a four-dimensional representation space; the V4 topography would presumably result from the mapping of a six-dimensional representation space and so would be accompanied by a more complex map. It is not clear that the currently available resolutions of optical imaging techniques would be able to detect such a potentially complex map.

CONCLUSION

The recursive filters map will need to be extended to account for other observed properties in V4. For instance, cells selective for textures and texture gradients have been isolated in V4 (Gallant et. al. 1993b). The recursive filter bank would be an ideal mechanism for analyzing textures, as many textures (such as the herringbone pattern) are defined not only by their distribution of energy in first-order features but also by the local spatial arrangement of these features.

Yoshioka et. al. (1992) have identified V4 cells that respond to complex stimuli derived from actual objects. The treatment of the representation space in this study is questionable, as only a few of the myriad possible patterns were explored. However, an enhanced representation space would assist the analysis of data such as this, and would lead to predictions of other possible patterns to test in single-unit experiments. The stimuli in this experiment varied in both the first-order feature space, and the local distribution of these features. It is not clear that all of these complex stimuli can be analyzed using the recursive filters model presented here, but it is possible that extending this model to allow for more levels of recursion may explain these results as well as results concerning response properties in IT.

REFERENCES

- Amir Y, Harel M, Malach R. (1993). Cortical Hierarchy Reflected in the Organization of Intrinsic Connections in Macaque Monkey Visual Cortex. *Journal of Comparative Neurology* 334: 19-46.
- Desimone R, Schein SJ. (1987). Visual Properties of Neurons in Area V4 of the Macaque: Sensitivity to Stimulus Form. *Journal of Neurophysiology* 57(3): 835-68.
- Desimone R, Schein SJ, Moran J, Underleider LG. (1985). Contour, Color and Shape Analysis Beyond the Striate Cortex. *Vision Research* 25(3): 441-52.
- DeYoe EA, Felleman DJ, Van Essen DC, McClendon E. (1994). Multiple Processing Streams in Occipitotemporal Visual Cortex. *Nature* 371: 151-4.
- Felleman DJ, Van Essen DC. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cerebral Cortex* 1: 1-47.
- Ferster D. (1987). Origin of orientation selective EPSPs in simple cells of the cat visual cortex. *Journal of Neuroscience* 7:1780-91.
- Gallant JL, Braun J, Van Essen DC. (1993a). Selectivity for Polar, Hyperbolic, and Cartesian Gratings in Macaque Visual Cortex. *Science* 259: 100-3.
- Gallant JL, Van Essen DC, Nothdurft HC. (1993b). Two-Dimensional and Three-Dimensional Texture Processing in Visual Cortex of the Macaque Monkey. In: *Early Vision and Beyond*. TV Papathomas, C Chubb, A Gorea, E Kowler eds. MIT Press: Cambridge, MA.
- Girard P, Salin PA, Bullier J. Visual Activity in Macaque Area V4 Depends on Area 17 Input. *Neuroreport* 2(2): 81-4.
- Heathcote A, Mewhort DJ. (1993). Representation and Selection of Relative Position. *Journal of Experimental Psychology: Human Perception and Performance*. 19(3): 488-516.
- Hubel DH, Wiesel TN. (1962). Receptive fields, binocular interaction and functional architecture of the cat's visual cortex. *Journal of Physiology (London)* 165:559-568.
- Kobatake E, Tanaka KJ. (1994). Neuronal Selectivities to Complex Object Features in the Ventral Visual Pathway of the Macaque Cerebral Cortex. *Journal of Neurophysiology* 71(3):856-867.
- Kritzer MF, Cowey A, Somogyi P. (1992). Patterns of Inter- and Intralaminar GABAergic Connections Distinguish Striate (V1) and Extrastriate (V2, V4) Visual Cortices and Their Functionally Specialized Subdivisions in the Rhesus Monkey. *Journal of Neuroscience* 12(11):4545-4564.
- Lund JS, Yoshlaka T, Levitt JB. Comparison of Intrinsic Connectivity in Different Areas of Macaque Monkey Cerebral Cortex. *Cerebral Cortex* 3(2): 148-62.

Somerville DC, Nelson SB, Sur M. (1995). An Emergent Model of Orientation Selectivity in Cat Visual Cortical Simple Cells. *Journal of Neuroscience* 15(8): 5448-65.

Yoshioka T, Levitt JB, Lund JS. (1992). Intrinsic Lattice Connections of Macaque Monkey Visual Cortical Area V4. *Journal of Neuroscience* 12(7): 2785-2802.

Figure 1. Two Patterns that can not be distinguished by first-order filters

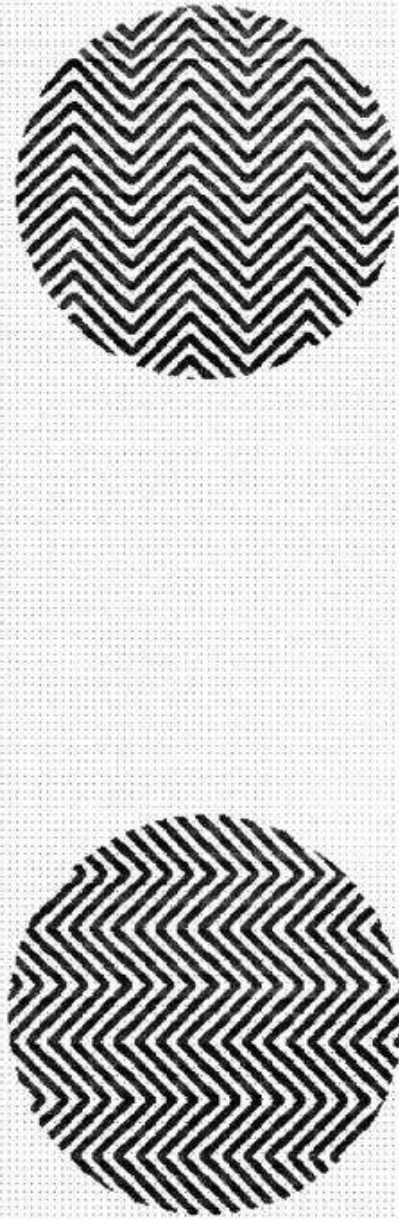


Figure 2. Analyzing the herringbone pattern

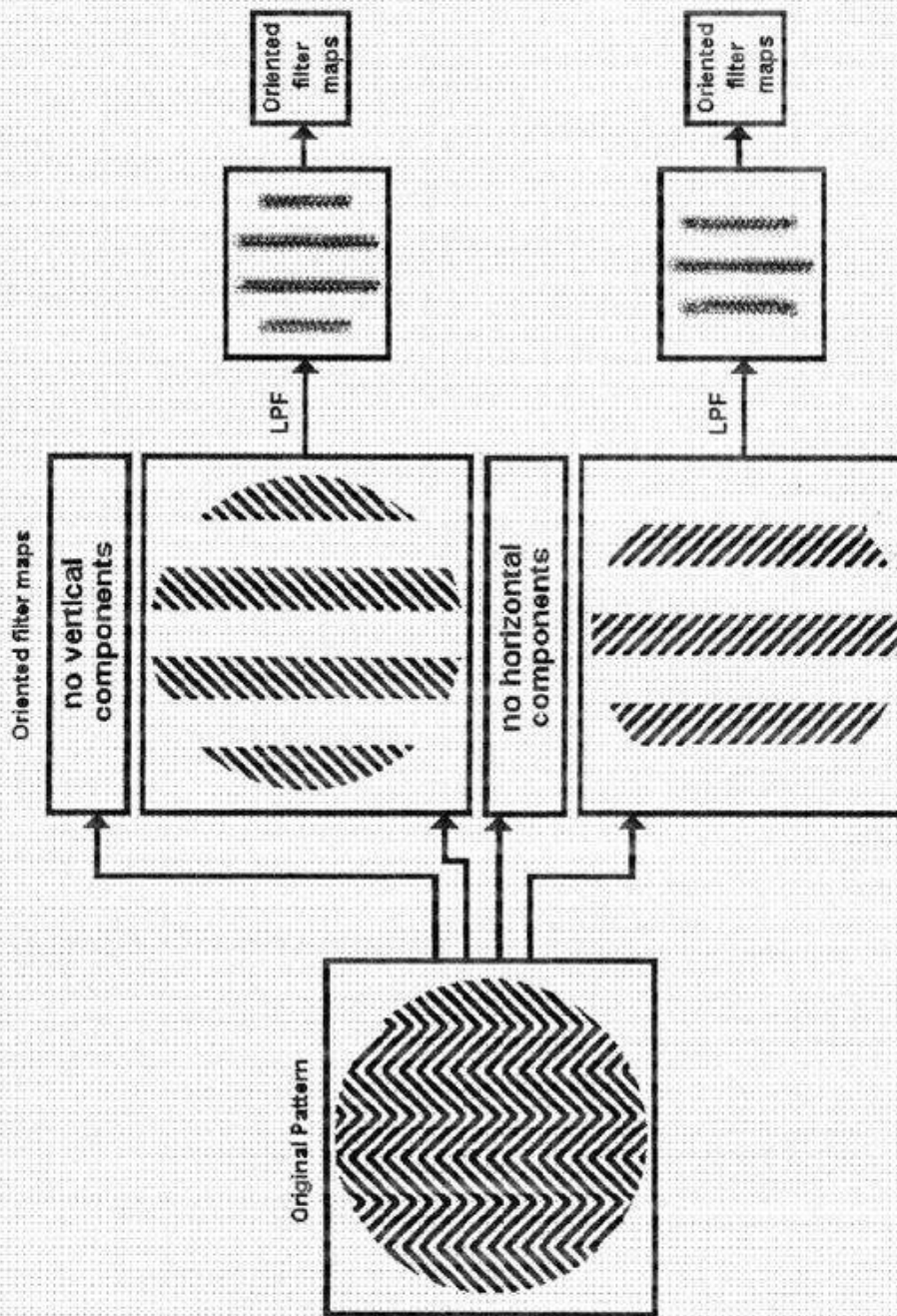


Figure 4. Recursive filtering of concentric stimulus

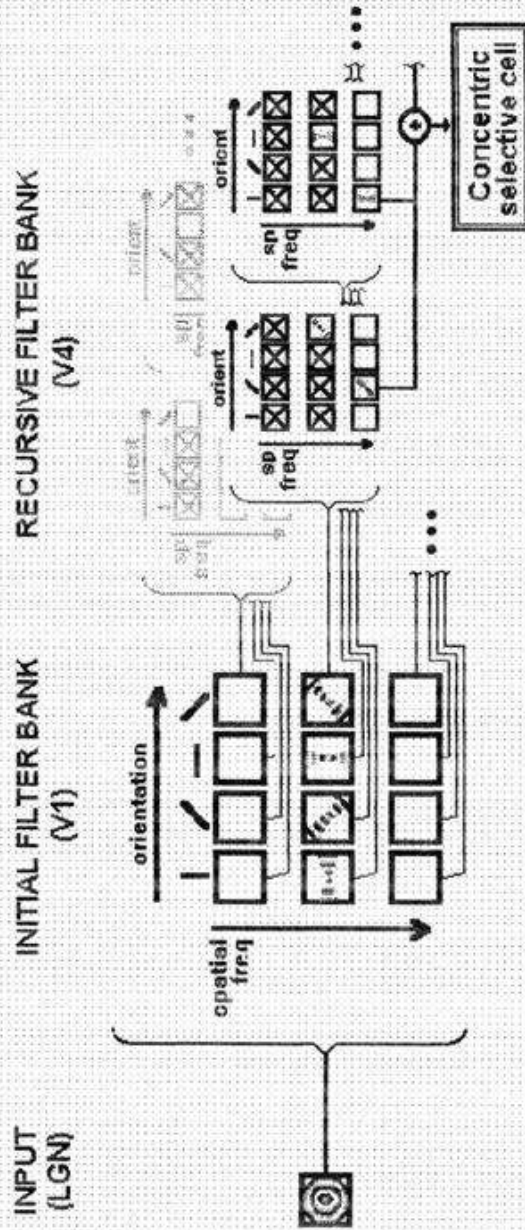
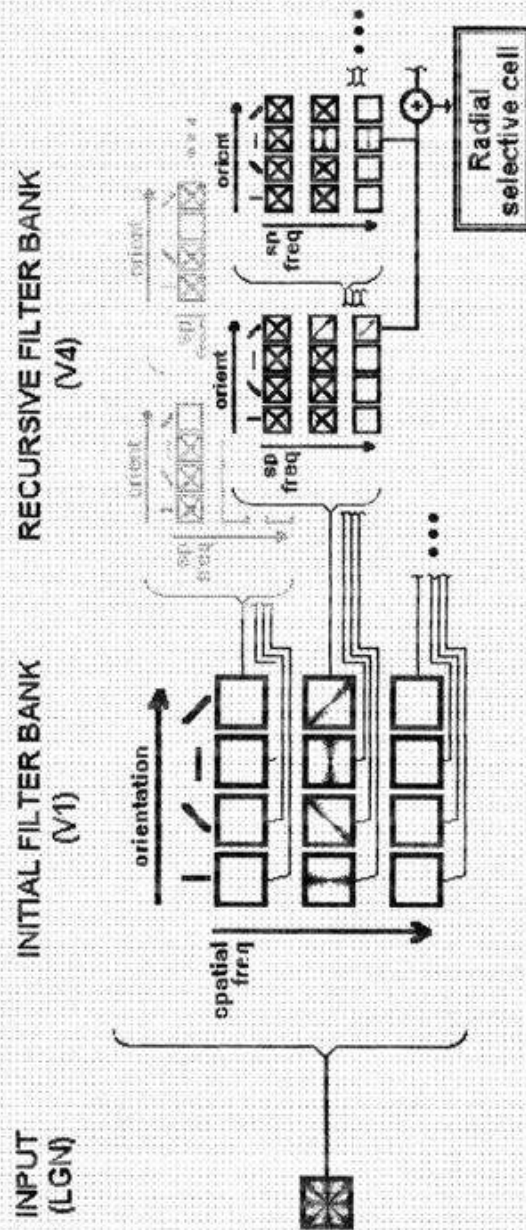
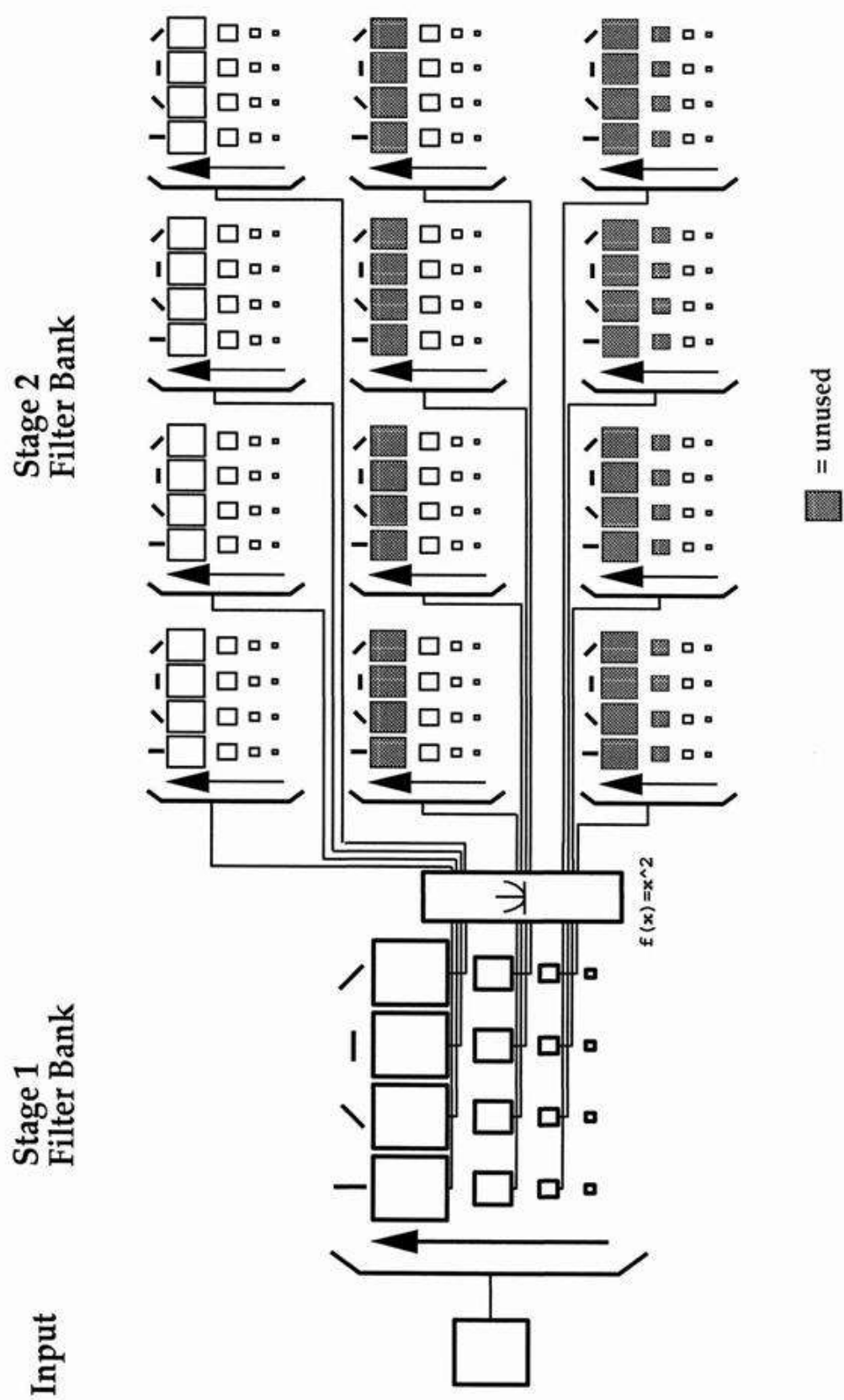
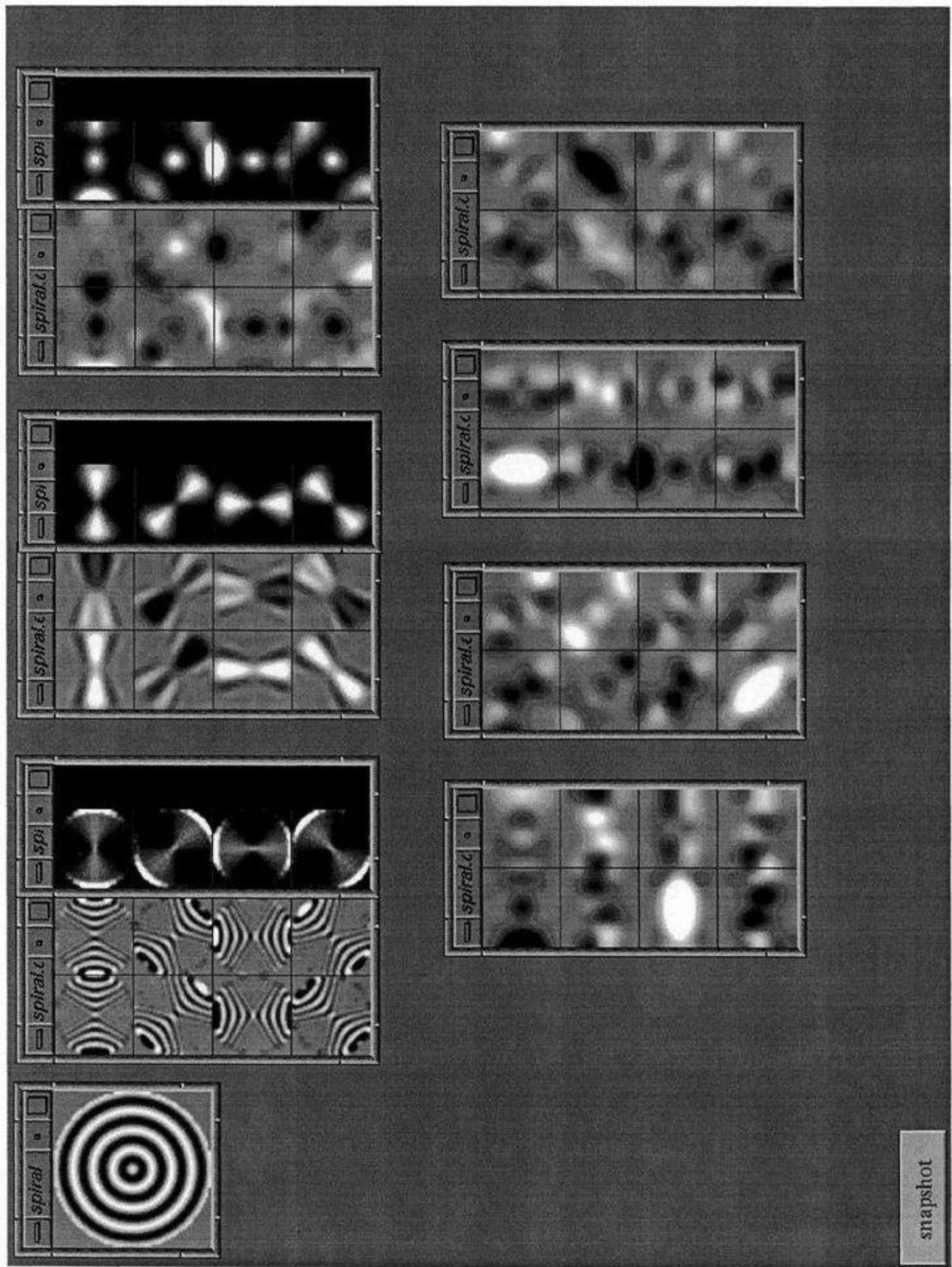


Figure 5. Recursive filtering of radial stimulus

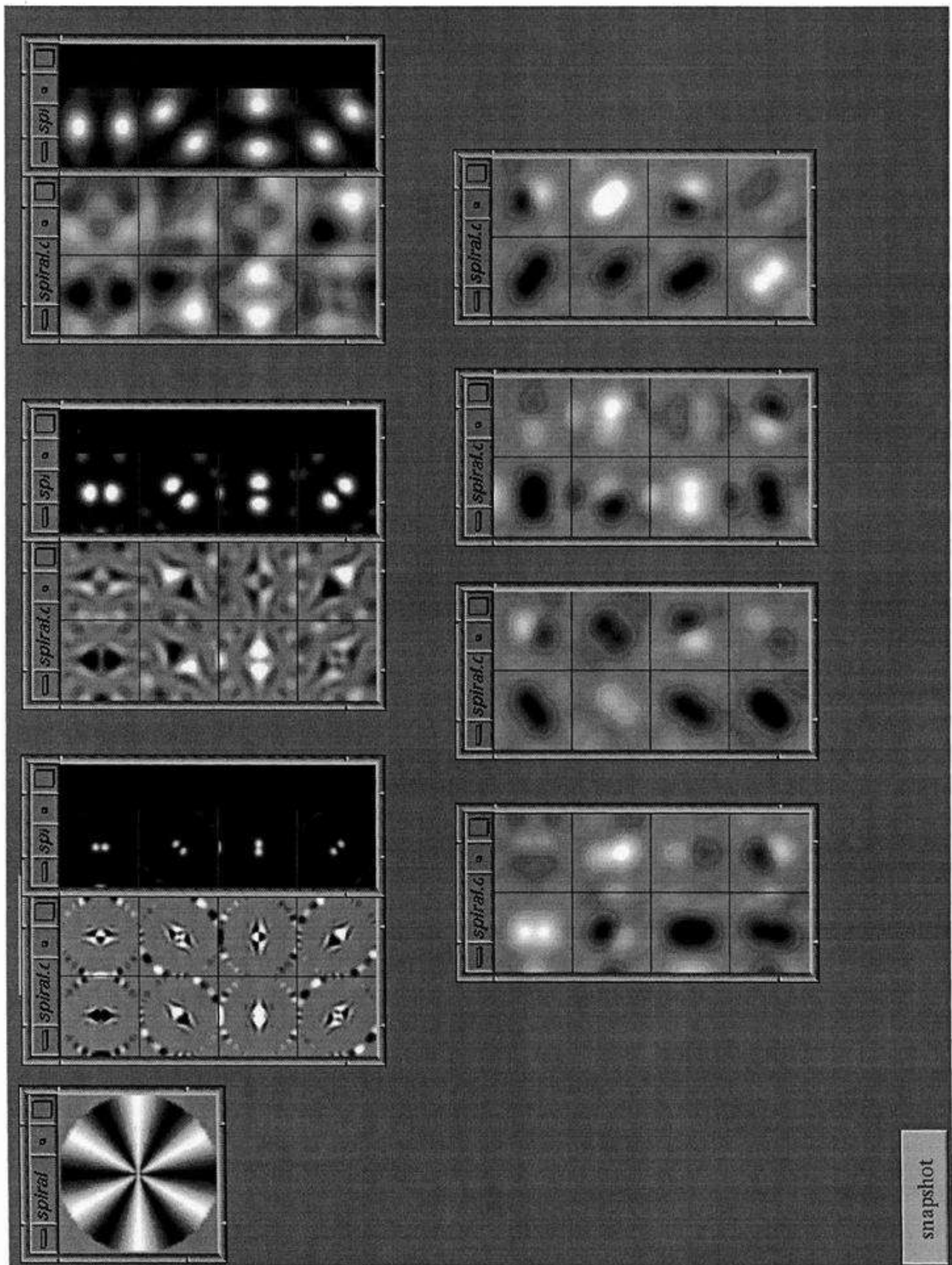


Recursive Filters Model





snapshot



Recursive Oriented Maps Paper

① LTG/NP

② Gallant Result for Lie operators

~~③ Gallant Result for~~

③ Recursive formulation of filters
for detecting Lie group stimuli
↳ why? sp. freq. → sp. freq.

④ Simulation results

↳ concentric

↳ radial

⑤ Extending to hyperbolic + spiral

⑥ Neurobiological basis of recursive filters

↳ theoretical economy

↳ models of orientation + spatial freq.
selectivity

↳ developmental economy