A Neural Network Architectural Model of Visual Cortical Cells for Texture Segregation

Giacomo M. Bisio, Daniele D. Caviglia, Giacomo Indiveri, Luigi Raffo, Silvio P. Sabatini
Department of Biophysical and Electronic Engineering
University of Genova-Via Opera Pia 11/A-16145 Genova-ITALY
ph: +39 10 3532163, fax: +39 10 3532795, e_mail: giacomo@cpsi7.dibe.unige.it

***Abstract*— In this paper we present a three layer hierarchical neural network architecture to be used in early vision processing tasks (e.g. texture segregation). Taking into account both the linear properties of simple cells receptive fields and the non-linear properties of intracortical processing, we defined the structure and the functionality of simple, complex and hypercomplex cells. The introduction in the model of hypercomplex cells which interact with complex cells, provides a complete feature extraction of textured images. Specifically, the first layer of the network extracts oriented textured elements, the second layer increases the sensitivity to texture differences and the last layer improves the selectivity of textural elements on the basis of their size.**

I. Introduction

In the mammalian visual cortex one can distinguish three types of neural cells: *simple, complex* and *hypercomplex*. Each of these cells is characterized by the type and the complexity of the luminous stimulus they are selective to: simple cells respond to oriented bars in specific conditions, i.e. position on the image plane and contrast direction (e.g. light bar on dark background or dark bar on light background); complex cells respond to oriented bars at a definite position but are insensitive to contrast direction, and hypercomplex cells respond to oriented bars or gratings of specific length, but are relatively insensitive to their position or contrast direction. Each type of cell, therefore, is characterized by a different *receptive field* which allows cortical neurons to act as precise and elaborate local feature extractors.

Many neurophysiologists [1][2] [3] have shown that simple cells have, for a broad range of inputs, a linear transfer function, while complex and hypercomplex cells have nonlinear characteristics. On the basis of such biological evidence we designed an artificial neural network that takes into account all three types of cells in order to model early vision processing and, in particular, texture segregation.

We exploited the linear behavior of simple cells and modeled them, as many other researchers did [4][5][6] [7], with a convolution operator, in which the convolution kernel represents the receptive field itself. In such a way, we were able to obtain the functionality of simple cells without having to consider all the neurons and their plentiful interconnections from the retina up to the visual cortex.

Subsequently, we modeled complex and hypercomplex cells introducing a hierarchical neural network, with local lateral connections and feedback pathways. Doing so, the model of the visual cortex results close enough to biological reality, yet simple enough for computer simulation and, eventually, hardware implementation in analog VLSI technology [8][9].

In the Section 2 we describe in detail both the model of simple cells and the structure of the network for complex and hypercomplex cells; in Section 3 we describe the behavior of the network and show some of the results obtained with textured test images; finally, in Section 4, we draw the final discussions and conclusions.

II. Description of the network

A. General overview

The model is a hierarchical multilayered network consisting of a sequence of artificial neurons strictly interconnected. Each neuron has a sigmoidal transfer function and excitatory synapses (only positive connection weights); complex and hypercomplex neurons, though, interact also with *inhibitory* neurons (negative connection weight). As in the real visual cortex, the synaptic strengths between different neurons (connection weights) are kept constant, therefore no learning takes place.

There are three main layers in the network, associated to the three types of cells considered. The input layer is the simple cell stratum, so we will refer to it as the simple layer. In a similar way we will call complex layer the layer of complex cells and hypercomplex layer the remaining one.

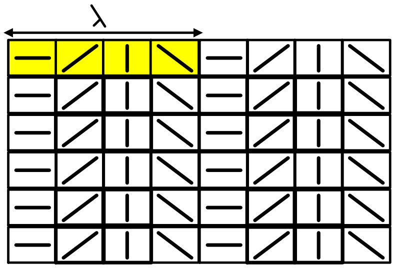


Fig. 1. Structural organization of orientation selective cortical neurons (cortical map). In dark background a hypercolumn of width λ is shown.

In each layer, cells are arranged in a highly functional way: neurons selective to a full set of different orientations are grouped together. Following the Hubel and Wiesel terminology [10] such structural entity is called a *hypercolumn*, while the structural arrangement of cells and hypercolumns on the cortical plane is called a *cortical map* (see Fig. 1). In our model all the cells of a hypercolumn are related to the same area on the image plane.

B. The simple layer

The simple layer is the first layer of the network. This layer is made of four types of simple cells. As anticipated in the introduction, simple cells are modeled functionally by a convolution operator, thus the excitation of a generic neuron with coordinates (u,v) on the simple plane is:

$$e(u,v) = \int\int I(x,y)\, w_{u,v}(x,y)\, dx\, dy$$

where $I(x,y)$ represents the input image and $w_{u,v}(x,y)$ the cell's receptive field.

The shape and the functionality of the receptive field of simple cells have been vastly investigated by numerous scientists [10] [11][12][13][14]. According to recent studies [15] , the mathematical function that best fits physiological data is a DOOG function, that can be written as:

$$\begin{aligned}\tilde{w}(x,y) &= A\,G(0,0,\sigma_x,\sigma_y) + B\,G(0,y_0,\sigma_x,\sigma_y) + \\ &\quad C\,G(0,-y_0,\sigma_x,\sigma_y)\end{aligned}$$

where

$$G(x_0,y_0,\sigma_x,\sigma_y) = exp\left\{-\left[\left(\frac{x-x_0}{\sigma_x}\right)^2 + \left(\frac{y-y_0}{\sigma_y}\right)^2\right]\right\}$$

and A, B, C, σ_x and σ_y are fitting parameters [7].

If $\sigma_x/\sigma_y >> 1$, the DOOG function models a receptive field selective to stimuli oriented along the x-axis, i.e. the *preferred orientation* of the receptive field is of 0°. Modifying σ_x and σ_y (while keeping constant their ratio) it's possible to change the dimensions of the receptive field and thus the specific spatial-frequency selectivity of simple cells.

The receptive field profile $w_{u,v}(x,y)$ varies in orientation preference θ and position (x_0,y_0) with the plane coordinates (u,v). Specifically, adopting the usual roto-translation equations, we can write:

$$w_{u,v}(x,y) = \tilde{w}(X(x,y,u,v), Y(x,y,u,v))$$

with

$$\begin{aligned}X(x,y,u,v) &= (x - x_0(u,v))\cos(\theta(u,v)) + \\ &\quad (y - y_0(u,v))\sin(\theta(u,v)) \\ Y(x,y,u,v) &= -(x - x_0(u,v))\sin(\theta(u,v)) + \\ &\quad (y - y_0(u,v))\cos(\theta(u,v))\end{aligned}$$

where (x_0,y_0) are the coordinates of the center of the receptive field and θ determines the direction of the preferred orientation.

By example, for the structural organization of the cortical map shown in Fig. 1, $\theta(u,v)$ results to be:

$$\theta(u,v) = \pi\,\frac{u}{\lambda}$$

where λ is the hypercolumn width.

In the present model, beside considering neurons selective to different orientations, we defined four different types of simple cells (see Fig. 2): ON-even cell (with function parameters A=2, B=C=-1), ON-odd cell (with function parameters A=0, B=-C=1), OFF-even cell and OFF-odd cell (with function parameters of opposite sign respect to the ON type cells). For all the cells defined we considered an aspect ratio r=3, i.e. $r = \sigma_x/\sigma_y = 3$

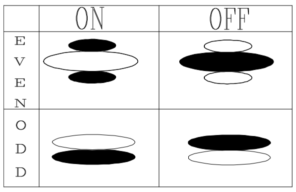


Fig. 2. ON, OFF, even and odd types of simple cells. The definition of ON-odd and of OFF-odd simple cells is arbitrary.

According to neurophysiological data [2] even cells lay beside odd cells (a shift of 90° in the receptive field's phase has been measured) and ON cells are near OFF cells; thus, taking into account the four types of simple cells and following Pollen and Ronner's suggestions, we defined a new cortical map, shown in Fig. 3. Four different simple cells (ON-even, OFF-even, ON-odd, OFF-odd), selective to the

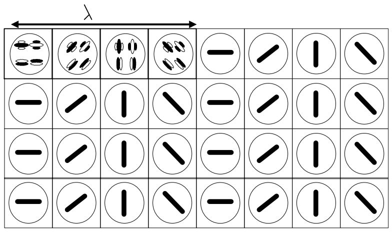


Fig. 3. Organization of the simple layer: this cortical map is equivalent to the one shown in Fig. 1 if we take into account the functional unit's output.

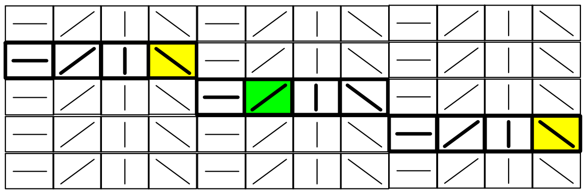


Fig. 4. Hypercolumns and complex cells involved in a cross orientation inhibition: the cell with orientation preference of 45° (in dark background) is inhibited by the two cells selective to 135° oriented bars (in gray background).

same orientation preference, are grouped together to form the basic functional unit of the simple layer array. The four cells are fully interconnected with inhibitory synapses to implement a *winner take all* network. Such new functional unit will respond to visual stimuli (as elongated structural elements) with the same orientation, that can differ either respect to phase or contrast. The output of such functional unit will provide the excitation to the succeeding layer.

C. The complex layer

The cortical map of the complex layer is the one shown in Fig. 1. Each complex cell is locally connected with cells belonging both to the simple plane and to the hypercomplex plane. Furthermore complex cells interact, in a recursive inhibitory way, with other cells belonging to the same layer. It has been shown, by numerous experiments [16] [17][18], that intralayer inhibitory connections play an important role in the visual cortex: receptive fields of cortical neurons are modified in order to obtain greater selectivity both in the frequency domain and in the spatial domain. Cortical inhibitory schemes are highly dependent on the structural topology of the cortical map and on the topology of the connections. The correspondence between *structural* inhibition schemes and *functional* inhibition schemes has been deeply studied by Worgotter et al., 1991. On the basis of these investigations, we implemented a *functional cross orientation inhibition* scheme: a cell with a θ degrees orientation preference is inhibited by cells belonging to neighboring hypercolumns (situated on an axis orthogonal to the cell's orientation preference) with orientation preference $\theta + 90°$ (see Fig. 4). As predicted by theory, the introduction of functional cross orientation inhibition enhanced system performance remarkably.

The interlayer connections, with the simple layer and the hypercomplex layer, are evidenced in Fig. 5: the complex cell considered receives feedforward input signals from three functional units (with an excitatory connection for the unit with same orientation preference and inhibitory connection for the two neighboring units with different orientation preference) and feedback input signals from three hypercomplex cells (with the same connection scheme as the one just described); moreover several complex cells with same orientation preference (but belonging to different hypercolumns) are connected with one cell of the succeeding layer.

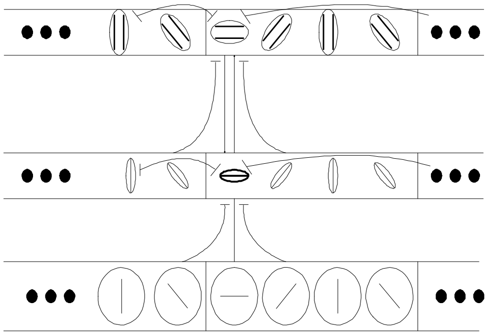


Fig. 5. Intralayer connection scheme. The hypercomplex cell, in the top row, receives excitation from the complex cell below, and from other complex cells here not shown, for the sake of simplicity.

D. The hypercomplex layer

In this layer, as in the preceding one, functional cross orientation inhibition is implemented, but the connection scheme with the complex layer is different, respect to the preceding one. As shown in Fig. 5, a hypercomplex cell receives input signals from different complex cells, and then returns a positive feedback signal to the complex neuron of the preceding layer with the same positional coordinates. The complex neurons that excite the hypercomplex one are all of the same orientation preference θ and belong to different hypercolumns located along a θ axis on the cortical map. A possible connection scheme of this type is shown in Fig. 6.

In the kind of network above described, complex and hypercomplex cells form a non-linear recursive system (cooperative loop), similar to others present in literature

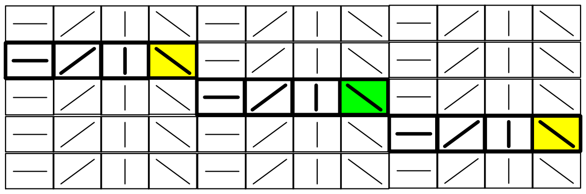


Fig. 6. Complex cells involved in the hypercomplex to complex feedback network: neurons of same orientation preference cooperate to excite a hypercomplex neuron (corresponding to the complex cell in dark background).

(e.g. the *boundary contour system* in Grossberg et al., 1989), that has the same functional characteristics as various ones observed in the human visual system, with psychophysical experiments [19]. From a strictly biological point of view, the topology of the network can be justified by the fact that there have been observed, in the visual cortex, long range connections [20][21] that, seemingly, join neurons with similar orientation preference.

Connecting relatively distant neurons, related to different (but nearby) areas of the image plane, it's possible to further enhance their specific selectivity. In fact, in such a way, the network detects different regions of the input image that have similar features, and, eventually, enhances the segregation of regions with different textural characteristics.

III. Application to texture segregation

To test our model we considered images of natural textures, taken from the Brodatz collection [22], commonly used in literature to evaluate the performance of these kind of systems.

The input image is subdivided in n regions, partially overlapped (n depends on the receptive field's dimension).

For each layer, taking the output cells with the same orientation, one generates four output images that encode the textural information extracted by this architecture. For example one image will show which neurons are selective to textural elements oriented along $0°$, the second to $45°$, the third to $90°$, and the fourth one to $135°$. Light regions of the output image indicate that the neurons are excited for the presence of the specific textural element while dark regions indicate that they are inhibited for the presence of textural elements of opposite type. Intermediate values of gray levels indicate that the neuron is silent (the input stimulus is nor optimal nor opposite).

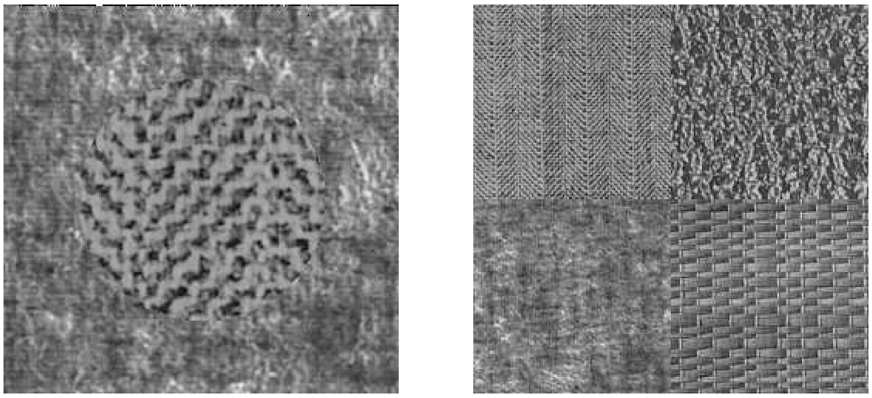
Examples of natural textures are shown in Fig. 7. The result of the first processing step (convolution and selection of the functional unit's output) on the image of Fig. 7.a, is shown in the top row of Fig. 8, while the result of functional cross orientation inhibition on the complex layer (without considering the feedback connections from hypercomplex neurons) is shown in middle row of Fig. 8 and the output of the overall system in the bottom row of the same figure. Examining the three outputs it is possible to evaluate, comparatively, the effects of the different operations done by the three layers: it is clear how the first stage of the network, which takes place on the simple layer, extracts features characterized by oriented structural elements, while in the second stage (i.e., the complex layer) the functional cross orientation scheme enhances the results obtained with the convolution by increasing image contrast, and in the final stage, on the hypercomplex layer, the cooperative loop, described in Section 2.4, improves spatial frequency selectivity and detects different patches of the image in which the structural elements extracted have same orientation.

The second test image, shown in Fig. 7.b is composed of four different textures, with different characteristics both in the spatial domain and in the frequential domain. For this image, we considered receptive fields with smaller dimensions and hence with sharper frequency selectivity. As shown in Fig. 9, the texture in the top right corner, with high frequency components in the vertical spatial axis, is well discriminated, while the two textures in the lower part of the input image, with low frequency characteristics, are not perfectly segregated.

IV. Discussion and conclusions

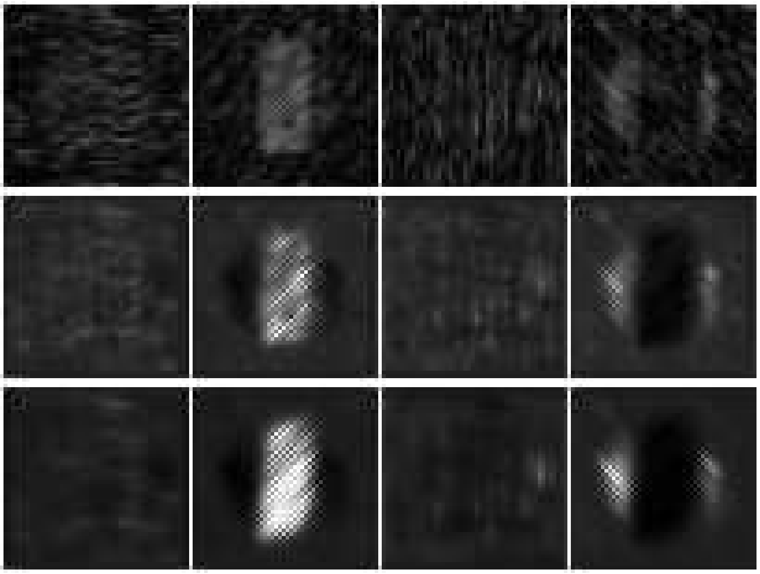
We conceived and implemented an artificial neural network with fixed connection weights (i.e. no learning is required) that models a few important tasks in early vision. The network has a hierarchical multilayered architecture that reflects the organization of the visual cortex in simple, complex and hypercomplex cells. We have shown how the network is able to discriminate textured images characterized by different oriented structural elements or different frequential properties. The network has an open architecture, thus it is possible to improve it's performances (better texture segregation) by considering more orientation selectivities (i.e., a hypercolumn with 8 neurons, instead of 4) or different frequency selectivities (i.e., receptive fields of different size). Work is being done to optimize frequency selectivity by considering several hierarchical networks, of the type described in this paper. The networks would differ in the size of simple cells receptive fields (hence in the dimensions of all the three layers) and would interact parallely.

The network here presented can also be considered as a neural feature extractor to be interfaced with more complex neural systems (which eventually implement unsupervised or supervised learning rules). In such a way, the overall system could pre-process the images to be analyzed, in order to reduce the dimensionality of the input



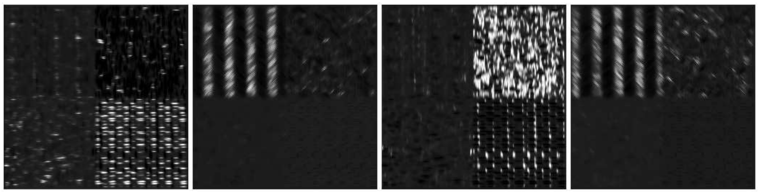
(a) (b)

Fig. 7. Example of textured images: in the first image (a) there are two different oriented textures in a uniform noisy background, while in the second one (b) there are four different textures disposed in quadrants.



$\theta = 0°$ $\theta = 45°$ $\theta = 90°$ $\theta = 135°$

Fig. 8. Output of the three different layers related to Fig. 7.a: in the simple layer (top row) structural elements are detected, but background noise is high. In the complex layer (middle row) image contrast is enhanced and textures are better distinguished. In the hypercomplex layer (bottom row) background noise is canceled and inhibitory regions appear clearer.



$\theta = 0°$ $\theta = 45°$ $\theta = 90°$ $\theta = 135°$

Fig. 9. Result of the hypercomplex layer elaboration on the textured image of Fig. 7.b: respect to Fig. 8 the receptive fields have smaller size and the number of neurons in the output layer has been increased correspondingly. In this case regions with high frequency components can be detected.

space and the computational cost and accomplish, at high speed and efficiency, more complex tasks, such as segmentation or classification. For example, our model could be used to perform *pre-attentive* vision tasks (contour or edge detection), and subsequent systems could perform *attentive* vision tasks.

Moreover, the particular topology of the network (hierarchical structure, local connections both intra layer and inter layers) and the absence of learning elements, suggests an analog VLSI implementation to substantially increase computational speed. An analog VLSI architecture of the network could be designed, with relatively little effort, from the start, or alternatively, part of it (specifically the complex and the hypercomplex layers) could be interfaced with other similar systems (already implemented in VLSI), that perform the same operations of the simple layer here described [23] [24].

References

[1] J. Kulikowski and P. Bishop, "Linear analysis of the responses of simple cells in the cat visual cortex," *Exp. Brain Res.*, vol. 44, pp. 386–400, 1981.

[2] D. Pollen and S. Ronner, "Visual cortical neurons as localized spatial frequency filters," *IEEE Trans. Syst. Man Cybern.*, vol. 13, pp. 907–916, 1983.

[3] R. Reid, R. Soodak, and R. Shapley, "Linear mechanisms of directional selectivity in simple cells of cat striate cortex," *Proc. Natl. Acad. Sci*, vol. 84, pp. 8740–8744, 1987.

[4] D. Marr, *Vision.* New York: Freeman, 1982.

[5] J. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Amer.*, vol. A/2, pp. 1160–1169, 1985.

[6] M. Turner, "Texture discrimination by Gabor functions," *Biol. Cybern.*, vol. 55, pp. 71–82, 1986.

[7] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," *J. Opt. Soc. Amer.*, vol. 7, pp. 923–932, 1990.

[8] C. Koch, "Seeing chips: Analog VLSI circuits for computer vision," *Neural Comp.*, vol. 1, pp. 184–200, 1989.

[9] L. Raffo, G. Bisio, D. Caviglia, G. Indiveri, and S. Sabatini, "A multi-layers analog VLSI architecture for texture analysis isomorphic to cortical cells in mammalian visual system," in *VLSI for Neural Networks and Artificial Intelligence* (J.G.Delgado-Frias and W.R.Moore, eds.), pp. 61–70, New York: Plenum, 1994.

[10] D. Hubel and T. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J. Physiol.*, vol. 160, pp. 106–154, 1962.

[11] V. Glezer, V. Ivanov, and T. Tscherbach, "Investigation of complex and hypercomplex receptive fields of visual cortex of the cat as spatial frequency filters," *Vision Res.*, vol. 13, pp. 1875–1904, 1973.

[12] K. Albus, "A quantitative study of the projection area of the central and the paracentral visual field in area 17 of the cat," *Exp. Brain Res.*, vol. 24, pp. 159–202, 1975.

[13] J. Robson, "Receptive fields. neural representations of spatial and intensive attributes of visual image," in *Handbook of Perception*, vol. 5, pp. 81–112, Academic Press, New York, 1975.

[14] R. Soodak, "Two dimensional modeling of visual receptive fields using gaussian subunits," *Proc. Natl. Acad. Sci*, vol. 83, pp. 9259–9263, 1986.

[15] M. Hawken and A. Parker, "Spatial receptive field organization in monkey V1 and its relatioship to the cone mosaic," in *Computationals Models of Visual Processing* (M. Landy and J. Movshon, eds.), pp. 83–93, MIT Press, 1991.

[16] A. Sillito, "Inhibitory mechanisms underlying the directional selectivity of simple, complex and hypercomplex cells in the cat's visual cortex," *J. Physiol.*, vol. 271, pp. 699–720, 1977.

[17] F. Worgotter and C. Koch, "A detailed model of the primary vision pathway in the cat: Comparison of afferrent excitatory and intracortical inhibitory connection schemes for orientation selectivity," *J. Neuroscience*, vol. 11, pp. 1959–1979, 1991.

[18] Y. Hata, T. Tsumoto, H. Sato, and K. Hagihara, "Inhibition contributes to orientation selectivity in visual cortex of cat," *Nature*, vol. 335, pp. 815–817, 1988.

[19] K. Prazdny, "On the nature of inducing forms generating perception of illusory contours," *Perception and Psycophysics*, vol. 37, pp. 237–242, 1985.

[20] C. Gilbert and T. Wiesel, "Clustered intrinsic connections in cat visual cortex," *J. Neuroscience*, vol. 3, pp. 1116–1133, 1983.

[21] D. Ts'o, C. Gilbert, and T. Wiesel, "Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis," *J. Neuroscience*, vol. 6, pp. 1160–1170, 1986.

[22] P. Brodatz, *Textures, a photographic album for artists and designers.* New York:Dover, 1966.

[23] C. Mead, *Analog VLSI and Neural Systems.* Addison-Wesley, Reading, 1989.

[24] J. Wyatt, C. Keast, M. Seidel, D. Standley, B. Horn, T. Knight, C. Sodini, H. Lee, and T. Poggio, "Analog VLSI systems for image acquisition and fast early vision processing," *Int. J. of Comp. Vision*, vol. 8, pp. 217–230, 1992.