Self-Organization in a Parametrically Coupled Logistic Map Network: A Model for Information Processing in the Visual Cortex

Ramin Pashaie, *Student Member, IEEE*, and Nabil H. Farhat, *Life Fellow, IEEE*

***Abstract*—In this paper, a new model seeking to emulate the way the visual cortex processes information and interacts with subcortical areas to produce higher level brain functions is described. We developed a macroscopic approach that incorporates salient attributes of the cortex based on combining tools of nonlinear dynamics, information theory, and the known organizational and anatomical features of cortex. Justifications for this approach and demonstration of its effectiveness are presented. We also demonstrate certain capabilities of this model in producing efficient sparse representations and providing the cortical computational maps.**

***Index Terms*—Chaos, cortex, cortical maps, information processing, self-organization.**

I. Introduction

In brain anatomy, the cerebral cortex (literally bark in Greek) is the outermost layer of the cerebrum and part of brain that is the center of unsupervised learning and the seat of higher level brain functions including perception, cognition, and learning of both static and dynamic sensory information. The cortex is composed of a sequence of discernable interconnected cortical networks (patches). Each concerned network has a specific functionality and is an ensemble of a large number of asymmetrically connected complex processing elements (CPEs) whose state–space orbits exhibit periodic orbits, as well as chaos and bifurcation. These networks of interconnected CPEs are responsible for the generation of sparse representations and efficient codes that are utilized during perception and recognition processes. Implementation of a microscopic model of cortex that incorporates small details of neurons, synapses, dendrites, axons, nonlinear dynamics of membrane patches, and ionic channels is prohibitively difficult even with the computational resources and the nanofabrication technologies available today or predicted for the future. Consequently, a realizable model should emulate such neurobiological computational mechanisms at an aggregate level.

As an example of cortical data manipulation, in this paper, we study and propose a model for cortical information processing in the self-organized visual system of mammals. Few other unsupervised learning models are attainable in the literature among which we address two outstanding ones. The first model is the one proposed by Field *et al.*, which is based on statistical structure of natural images to produce efficient codes [1]–[4]. They showed that a coding strategy that maximizes sparseness is sufficient to develop a full set of localized, oriented, bandpass receptive fields that span the image space, similar to those found in the visual cortex. The other model, the so-called laterally interconnected synergetically self-organizing map (LISSOM), and related models were introduced by Miikkulainen *et al.* [6]. The LISSOM is similar to the well-known Kohonen's self-organizing map [8], which is modified by adding lateral connections. This model simulates the structure, development, and function of the visual cortex at the level of cortical maps and their connections. Computational units of LISSOM are cortical columns that continuously adapt to afferent and lateral inputs, and the units synchronize and desynchronize their activity. This model has been partially successful in interpreting neurobiological facts such as columnar map organization as well as patchy connectivity, recovery from retinal and cortical injury, psychophysical phenomena such as tilt after effect [7], contour integration, and preference for faces [5].

In this paper, based on the tools of nonlinear dynamics and information theory, a set of equations that mimics the way cortex processes information and interacts with subcortical areas is described. We model the cortical networks with networks of parametrically coupled nonlinear iterative maps each having complex dynamics that represents populations of randomly interconnected neurons possessing collective emergent properties. This biologically inspired model incorporates several known or plausible organizational attributes of the cortex. The first model of this type was introduced and investigated earlier by Farhat [9]–[11].

A stylized model of the visual system of mammals is illustrated in Fig. 1(a). In these species, early image processing begins in the retina and continues in the lateral geniculate nucleus (LGN) by taking the spatial–temporal derivative of the image and reducing the redundancy of the information in part. Then, the LGN sends projections to the primary visual cortex, area V1.

Manuscript received August 31, 2007; revised February 02, 2008, February 02, 2008, and February 02, 2008; accepted November 27, 2008. First published March 06, 2009; current version published April 03, 2009. This work was supported in part by Army Research Office MURI under Grant DAAD 19-01-0603 via Georgia Institute of Technology Subcontract E-18-677-64, in part by the U.S. Office of Naval Research under Grant N00014-94-1-0931, and in part by Army Research Office DURIP instrumentation under Grant W911NF-04-1-0177.

R. Pashaie is with the Bioengineering Department, Stanford University, Stanford, CA 94305 USA (e-mail: raminp@stanford.edu).

N. H. Farhat is with the Electrical and System Engineering Department and the Mahoney Institute of Neurological Sciences, University of Pennsylvania, Philadelphia, PA 19104 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNN.2008.2010703

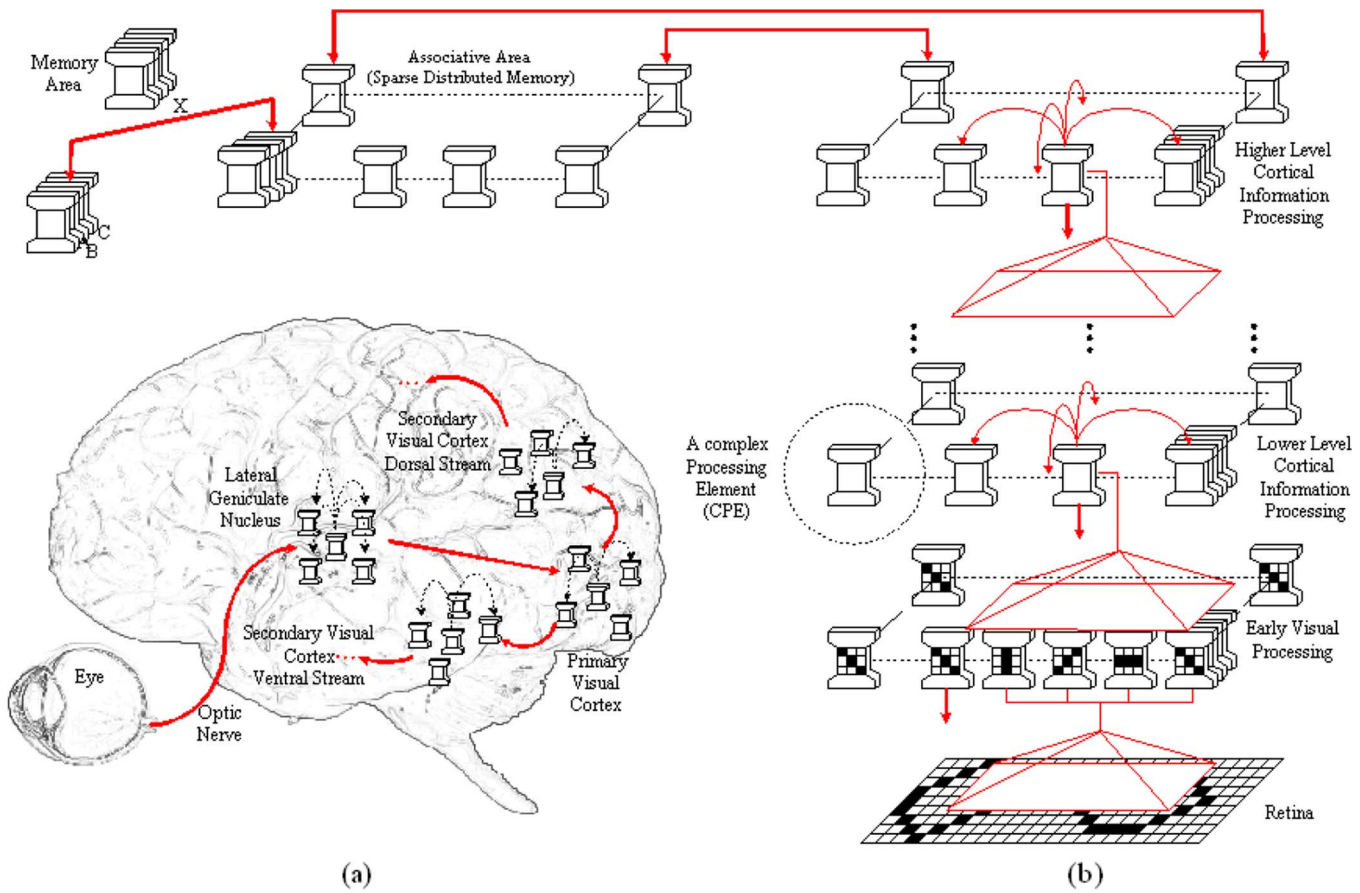

Fig. 1. (a) Flow of information in the visual system of mammals. Information about the image that is projected on and preprocessed by the retina propagates along the optic nerve until LGN. The LGN, which is part of the thalamus, continues the preprocessing and feeds the visual cortex. In the visual cortex, data is processed by a sequence of cortical networks. (b) Organization of the cortical processing model. This network is composed of multiple layers of interconnected complex processing units. In biology, there are also feedbacks from upper layers to the lower layers. The role of these feedbacks (that could be significant) are not well understood and are not considered in this minimal model. Basics of the block diagram shown in Fig. 1(b) are adapted from [12]; however, the computational units and architecture of the network presented here is quite different from [12].

The output of V1 flows through the dorsal and ventral pathways in the secondary visual cortex to extract concepts of where/what that are associated with the image. This modality is similar to the multilayered network illustrated in Fig. 1(b). Here a sequence of cortical layers are fed by a preprocessed image developed in the early stages. Then, each cortical layer increases the efficiency and sparsity of the representation by reducing the redundancy and extracting the critical features of the image that play the vital role in cognition. Finally, the most efficient code is projected to the associative memory that can be realized as a sparse distributed associative memory [13]. The output of the associative area addresses a memory area where the target representations are stored. Study of the dynamics of the cortical patches in this model is the main focus of this paper.

II. Architecture of the Network of Parametrically Coupled Complex Processing Elements (PCLMN)

In general, the architecture of a network is clarified if we address two questions: What is the mathematical model that governs the dynamics of each individual processing unit in the network? How these processing units are coupled? Regarding our cortical model, we answer these two questions in the following sections.

A. Netlets: Computational Units of Cortex

Netlets are sets of discrete populations of randomly interconnected neurons [14], [15]. The expected fraction of active neurons in a successive interval $<\alpha(n+1)>$ as a function of the fraction of active neurons in the previous time interval $\alpha(n)$ can be expressed in a recursive format: $<\alpha(n+1)>=\Psi(\alpha(n))$. In this scheme, Anninos *et al.* [15] has shown that for a single, isolated probabilistic net with excitatory and inhibitory synapses the expectation value $<\alpha(n+1)>$ can be approximated as follows:

$$\langle\alpha(n+1)\rangle \simeq [1-\alpha(n)]e^{-\alpha(n)h\mu^-}\sum_{m=0}^{M}\frac{[\alpha(n)h\mu^-]^m}{m!} \times\left\{1-e^{-\alpha(n)(1-h)\mu^+}\sum_{l=0}^{\eta-1}\frac{[\alpha_n(1-h)\mu^+]^l}{l!}\right\}. \quad (1)$$

In this model, h represents the fraction of inhibitory neurons in a netlet, $\mu^+(\mu^-)$ is the average number of neurons in a netlet with afferent connections from a given excitatory (inhibitory) neuron in the netlet, η is the minimum number of excitatory and

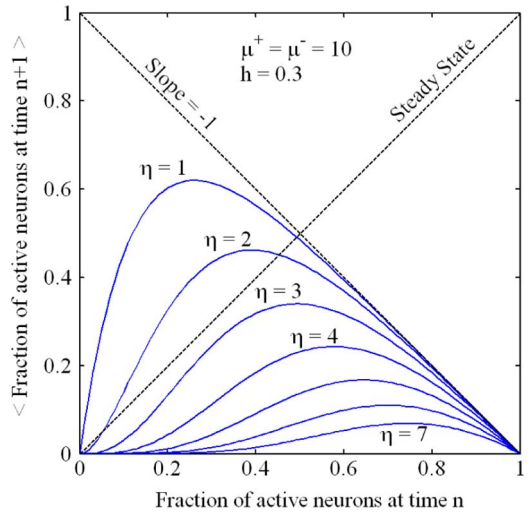

Fig. 2. Fraction of active nodes of a netlet, with the same amount of excitatory and inhibitory connections, at the moment $n+1$ as a function of active nodes at the previous moment. Different curves correspond to different numbers of presynaptic spikes that are necessary to elicit a postsynaptic spike.

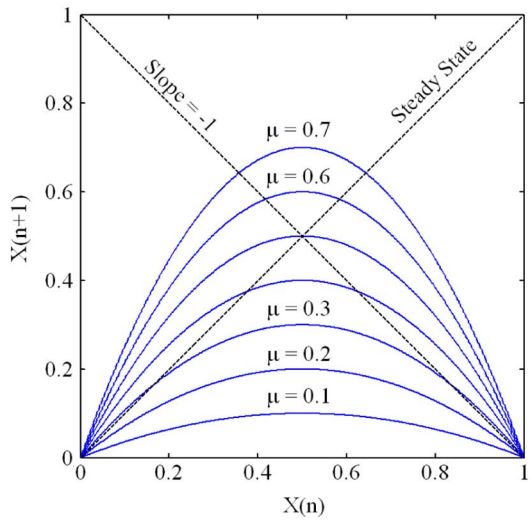

Fig. 3. Group of quadratic functions employed in the generation of the logistic map. The curves in this figure are quadratic functions that are plotted for different values assigned to the bifurcation parameter $\mu \in [0, 1]$.

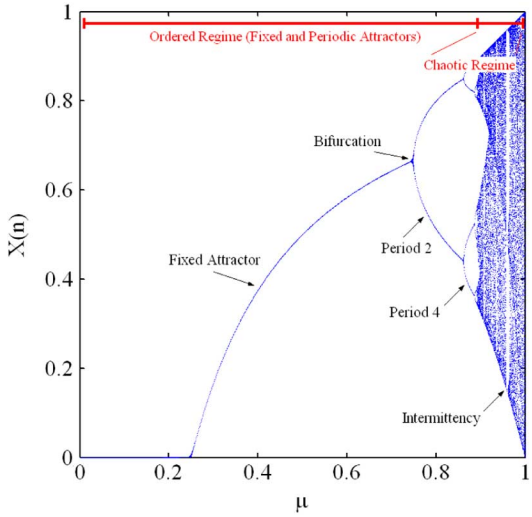

Fig. 4. Bifurcation diagram of a logistic map. Dynamics of this CPE changes from fixed attractor to bifurcation and chaos when the bifurcation parameter increases in the allowed region.

inhibitory inputs necessary to trigger a neuron, which has received m inhibitory inputs, and M is the total number of inhibitory connections. Fig. 2 shows the graphs of $< \alpha(n+1) >$ versus $\alpha(n)$ for a netlet with the same amount of excitatory and inhibitory connections $\mu^+ = \mu^- = 10$ and $h = 0.3$.

Plots of $< \alpha(n+1) >$ versus $\alpha(n)$ obtained under a range of circumstances and assumptions are found to invariably resemble a distorted version of the quadratic or logistic map (Fig. 3), a nonlinear iterative map on the unit interval that exhibits complex orbits and bifurcation between them depending on the value of a nonlinearity parameter. The similarity between the netlet's return map and that of logistic map has been noted by Harth [16] who also mentions that complex and unpredictable sequences were observed in some of their early simulations of the netlets at the time, suggesting to him that certain regions of the netlet's parameter space can lead to observation of chaos in addition to the periodic and fixed point modalities they usually observed. In the light of this evidence, the intriguing conjecture is that the cortical networks can be modeled in an efficient way by means of populations of logistic processing elements. We should keep in mind that a net of N bifurcation processing elements represents (10^2 to 10^3) $\times N$ cortical neurons. As a result, our network emulates the cortical computational mechanisms at an aggregate level. Other types of map-based computational models have been employed previously in the simulation of large scale cortical networks [20]. Nevertheless, our aggregate model, in addition to its simplicity, introduces a new level of abstraction.

As we discussed, dynamics of each CPE in a cortical patch is modeled by a logistic map. Hence, in a 2-D network of $N \times N$ computational units, the activity of the ijth CPE at any successive interval $X_{ij}(n+1)$ can be expressed as a function of its activity at the previous interval $X_{ij}(n)$

$$X_{ij}(n+1) = 4\mu_{ij}(n)X_{ij}(n)(1 - X_{ij}(n)) \qquad (2)$$

where $i, j = 1, 2, \ldots, N, X_{ij} \in [0, 1]$, and $\mu_{ij} \in [0, 1]$ represents the bifurcation parameter of the ijth nonlinear map. Evolution of the logistic map, expressed by (2), depends crucially on the value of its bifurcation parameter μ_{ij}. Depending on the value assigned to μ_{ij} in the allowed interval, the asymptotic possible states of the return map include fixed-point, period-m, and chaotic attractors. The bifurcation diagram of a logistic map where the asymptotic solution is plotted as a function of the bifurcation parameter is shown in Fig. 4. We should also mention that a variety of different enabling technologies, namely, electronics and photonics, have been studied for realization of large arrays of such nonlinear maps [17]–[19].

B. Coupling of the Complex Processing Elements

Next step in constructing the network is to couple the processing elements based on the biological observations. In this regard, Freeman's experiment is inspiring [21], [22]. Freeman *et al.* who were recording electroencephalogram (EEG) signals of the olfactory bulb of rabbits, experimentally proved that the

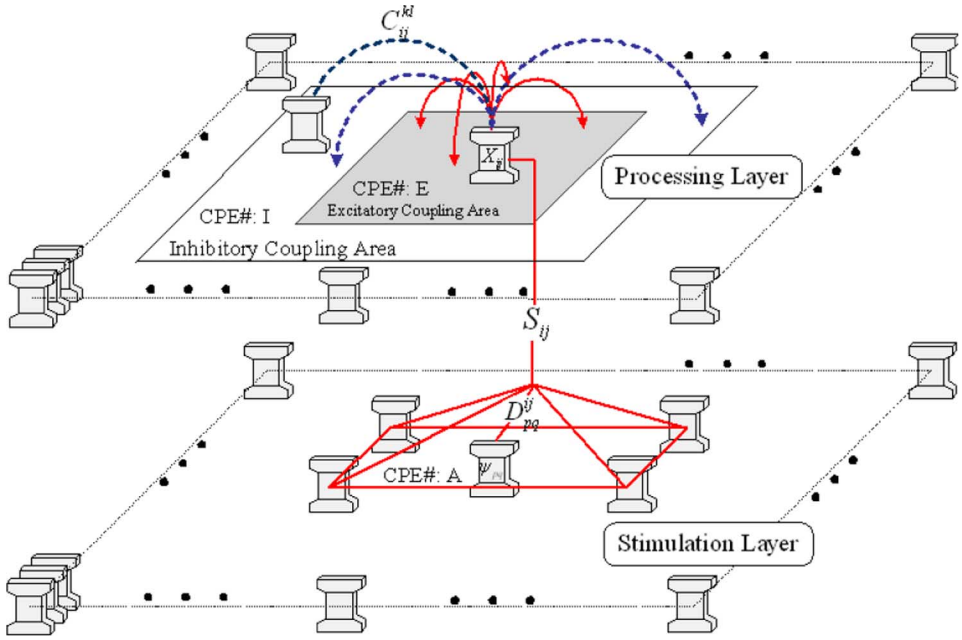

Fig. 5. Coupling of two cortical patches. The ijth CPE in the processing layer is stimulated by "A" number of processing elements in the stimulation layer. The ijth element is also excitatorily coupled to the "E" number of nearest neighbor CPEs and inhibitorily coupled to the "I" number of far neighbor CPEs. D_{pq}^{ij} is the afferent coupling coefficient from the pqth stimulating CPE to the ijth processing CPE and C_{ij}^{kl} is the lateral coupling coefficient from the ijth to klth CPEs in the processing layer. S_{ij} is the accumulated afferent input from the stimulation layer to the ijth processing CPE.

nervous activity of the olfactory system remains chaotic when a rabbit smells an unfamiliar odor, indicating lack of cognition, but switches quickly from a chaotic state to a periodic state when a familiar odor, which was already learned by the rabbit, is presented. Freeman assumed that the dynamics of the rest of brain is similar to the olfactory bulb and he suggested that chaotic activities in brain serve as the ground state of perception process, source of intelligence, consciousness, etc. Experimental observations and theoretical models endorse Freeman's suggestions [23], [24]. Chaos can also provide a mechanism to generate controllable noise that can push the state of the network out of spurious attractors [25], [26].

From the Freeman's experimental observations, it can be induced that if a CPE is trained to be sensitive to a particular feature in a stimulus, once such a stimulation is presented, the dynamics of this CPE will change from chaotic regime to periodic or fixed attractors (ordered regimes). In the network of logistic maps, behavior of the CPEs is under control by adjusting the bifurcation parameters so that a familiar aspect in the afferent or lateral stimulation changes the asymptotic state of the processing element by reducing its bifurcation parameter. At the same time, an unfamiliar stimulation pushes the dynamics of the CPE to the chaotic regime by increasing the bifurcation parameter. Hence, the processing elements of this network are coupled parametrically and the network is called parametrically coupled logistic map network (PCLMN). The explicit mathematical expressions of parametrically coupling are given in the next section.

The pioneer research in the study of the dynamics of the coupled map lattices (CMLs) was performed by Kaneko [30], [31]; however, the PCLMN we study in this paper distinguishes itself from Kaneko's CML because of the following: 1) the couplings in the PCLMN are nonlinear and 2) unlike diffused couplings in the CMLs, in the PCLMN, the couplings are indirectly applied through the bifurcation parameters. The idea of parametric coupling has been previously adapted to improve the performance of the associative memories such as Hopfield model [27], [28] and biomorphic dynamical networks that are used for cognition and control purposes [29], [9].

III. Dynamics of the Network

We explain the dynamics of the network by studying a pair of coupled PCLMNs illustrated in Fig. 5. Here the lower layer (stimulation layer), which could be an early processing network or a cortical patch, stimulates the upper layer (processing layer). The activity of the ijth processing element in the processing layer is expressed by X_{ij} whereas ψ_{pq} represents the activity of the pqth processing element in the stimulation layer. When the stimulation layer is a cortical patch, ψ_{pq} is the normalized entropy of the activity of the pqth processing element. However, ψ_{pq} represents the normalized stimulation intensity on the pqth CPE when the stimulation layer is an early processing network. Accordingly, $X_{ij}, \psi_{pq} \in [0,1], \forall i,j,p,q = 1,\ldots,N \cdot D_{pq}^{ij} \in [0,\infty)$ is the afferent coupling coefficient from the pqth CPE in the stimulation layer to the ijth CPE in the processing layer and $S_{ij} \in [0,1]$ is the accumulated (summed) afferent input to the ijth CPE. $C_{ij}^{kl} \in [0,\infty)$ is the lateral coupling coefficient of the ijth CPE to the klth CPE in the processing layer. In this figure, "A" is the cardinality of the set of input elements that are coupled to any CPE at the upper layer. An essential tool for sparsification is to have both excitatory and inhibitory lateral couplings. Therefore, in the processing layer, each CPE is excitatorily coupled to the "E" number of nearest neighbors and inhibitorily coupled to the "I" number of far neighbor CPEs. This

type of coupling produces local cooperation and global competition among CPEs of a cortical patch that ultimately plays a vital role in the generation of efficient representations [6]. The dynamics of this network has these three different stages: stimulation, iterations-updates, and adaptation.

A. Stimulation

We start the process with an undeveloped network (tabula rasa), by which we mean that all the afferent and lateral coupling coefficients D_{pq}^{ij} and C_{ij}^{kl}, respectively, are set to 1.0 and all the processing units are in full chaotic regime $\mu_{ij}(0) = 1.0$. Then, an $N \times N$ afferent input stimulates the processing layer

$$\mu_{ij}(m=1) = \sigma(1 - \xi S_{ij}) \tag{3}$$

$$S_{ij} = \frac{1}{A} \sum_{pq} \psi_{pq}^{D_{pq}^{ij}} \tag{4}$$

$$\sigma(x) = \begin{cases} 1, & x \geq 1 \\ x, & 0 < x < 1 \\ 0, & x \leq 0. \end{cases} \tag{5}$$

In (5), $\sigma(x)$ is defined as a piecewise linear sigmoidal function and ξ is a parameter that controls the effectiveness of the afferent stimuli (inputs from stimulation layer to the CPEs of the processing layer) compared to the lateral inputs (inputs from other CPEs in the processing layer). m in (3) is an integer that counts the number of updates (which will be clarified further in Section III-B). If the stimulation from the lower layer does not affect the ijth processing unit of the upper layer $S_{ij} = 0$, its bifurcation parameter μ_{ij} remains 1.0. However, any $S_{ij} > 0$ reduces the bifurcation parameter and a strong excitation decreases the bifurcation parameter to lower values in the periodic or even fixed-point attractor regimes.

As mentioned earlier, in an undeveloped network, all the afferent coupling factors are set to 1.0 so that the stimulations (3) and (4) work as a blurring filter without any specific preference; however, later we will see that the adaptation process modifies the coupling coefficients in a way that (3) and (4) in a developed network behave as localized, orientation selective, spatially bandpass filters.

B. Iterations Updates

In our model, the evolution of the activity of each CPE is a function of the pattern of activity of that CPE and the surrounding CPEs. We use the entropy of the activity of a CPE and the joint entropy between the activities of a CPE and those of the surrounding CPEs as a measure to quantify different patterns of activities. For instance, a chaotic pattern of activity has bigger entropy compared to periodic activities and fixed-point attractor regime, which has zero entropy. Plots of the entropy of asymptotic values of the dynamics of a logistic map and the normalized joint entropy between patterns of activities of two logistic maps as a function of the bifurcation parameters are displayed in Fig. 6. What is important in the current model is not the level of activity of a CPE but its pattern of activity which could change from ordered patterns to chaos. Ordered patterns have lower entropy compared to chaotic patterns.

In this model, a CPE with a fixed or periodic activity affects other CPEs more efficiently compared to a CPE with chaotic

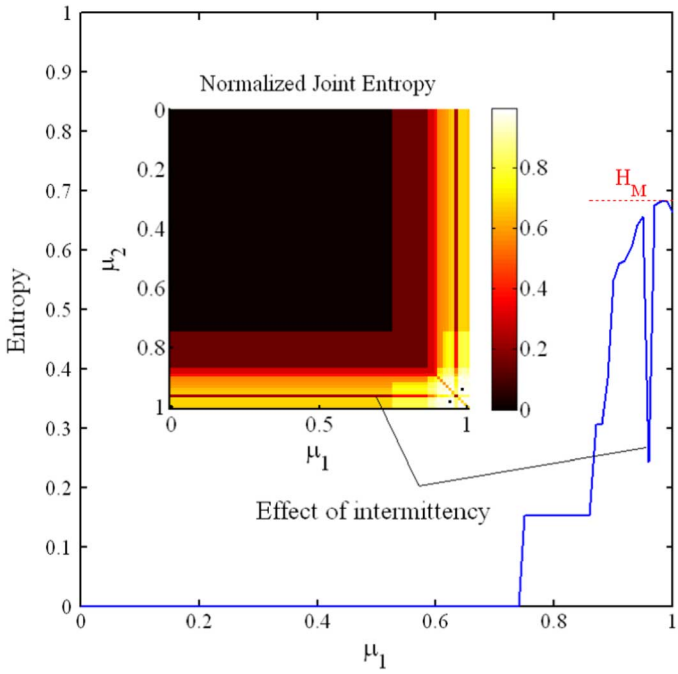

Fig. 6. Entropy and normalized joint entropy of logistic maps plotted as a function of the bifurcation parameters. The sharp drops occur as a result of the intermittency in the dynamics of the logistic maps. The joint entropy of two logistic maps is reminiscent of the well-known Hebbian learning rule (see Section III-B for more details).

activity. This statement is reflected in the mathematical expression of the parametrical coupling

$$\mu_{ij}(m) = \sigma\left\{[1 - \xi S_{ij}] + \left[-\frac{\gamma}{E}\sum_{k,l}\left(1 - \frac{H_{kl}}{H_M}\right)^{C_{ij}^{kl}}\right] + \left[\frac{\lambda}{I}\sum_{k,l}\left(1 - \frac{H_{kl}}{H_M}\right)^{C_{ij}^{kl}}\right]\right\}. \tag{6}$$

In (6), H_{kl} is the entropy of the klth CPE, and H_M is the maximum possible value of the logistic map's entropy (Fig. 6). The parameter C_{ij}^{kl} represents the coupling coefficient of the ijth to the klth CPEs in the processing layer. As (6) implies, the bifurcation parameter of the ijth CPE decreases as a result of ordered activities of the nearest neighbors. On the other hand, ordered activities of the far neighbors increases the bifurcation parameter of the CPE. At the same time, CPEs with chaotic activities (higher entropies) do not have a considerable effect on the surrounding CPEs.

The parameters γ and λ in (6) are the coefficients that control the effectiveness of the excitatory and inhibitory lateral connections relative to the afferent stimulation. After stimulating the network, using (3)–(5), we *iterate* all the nonlinear maps by (2). Then, after few hundreds of iterations, we compute the entropy of each CPE based on its activity during iterations and we *update* all the bifurcation parameters by employing (6). This protocol is repeated until the convergence of the population kurtosis of the activities of the CPEs in the processing layer (population kurtosis and its properties will be explained in the next

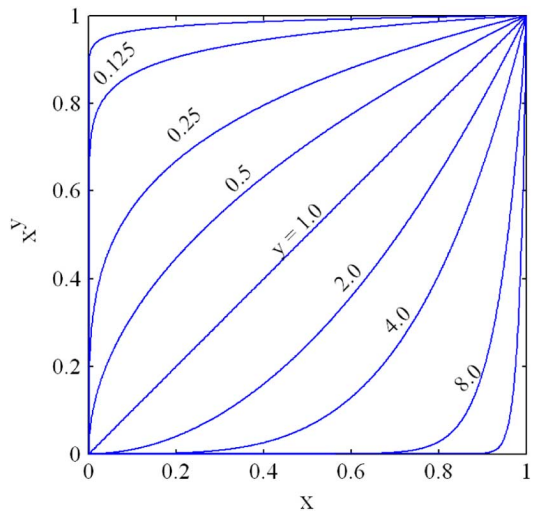

Fig. 7. Stimulation of a cortical patch with the output of an early processing layer. In this example, the early processing layer functions simply as an edge enhancing filter. (a) The image that is fed to the early processing layer (left) and the output that goes to the processing layer (right). (b) At the beginning, for all is and js, the coupling factors are $D_{pq}^{ij} = C_{ij}^{kl} = 1.0$. Other parameters that are involved in the model are $\xi = 0.18, \gamma = 0.15, \lambda = 0.18, \eta = 0.1, \beta = 1.0$, and $A = 1, E = 81, I = 360$. The set of images and the corresponding histograms (below each image) show the evolution of the normalized entropies (horizontal axis in the histograms) of the activity of CPEs in the cortical patch. (c) and (d) Iterations and updates of the cortical network after 1-step and 2-steps adaptation. (e) Effect of the adaptation on afferent coupling coefficients D_{pq}^{ij}s. The very small dark spots scattered around in the images are caused by intermittency in the dynamics of logistic maps.

section). The variable m in (3) and (6), counts the number of updates and the variable n in (2) is the number of iterations. The value of n is reset after each update. In our simulations, these two parameters smoothly change as a function of time: γ decreases, $\gamma(m) = \gamma_0 + 0.1\gamma_0 \exp(-m/10.0)$, and λ increases, $\lambda(m) = \lambda_0 + 0.2\lambda_0(1 - \exp(-m/10.0))$. As a result, at the beginning, higher values of γ push more numbers of CPEs toward ordered activity regime and bring out more details of the stimulation that are required for cognition. When the initial moments are passed, an increase in the value of λ improves the capability of the CPEs to produce more sparse representations.

Another point is that in (6) we have applied the coupling through functions that look like x^y where x and y are real numbers and $x \in [0, 1], y \in [0, \infty)$. Curves of this function for different values of y are displayed in Fig. 7. Depending on the range of values of y, the function x^y has three different phases. When $y = 1.0$, the coupling is linear. Otherwise, the coupling is nonlinear for $y < 1.0$ and $y > 1.0$. For $y \ll 1.0$, the value of x^y is big and the coupling is strong even when the input is weak. On the other hand, for $y \gg 1.0$, the value of the function x^y is small and the coupling is weak unless the input is very strong.

C. Adaptation

After few updates of the bifurcation parameters, it is the time for the network to train (self-organize) itself by modification of the coupling factors. For training, we use the following couple of substitution relations:

$$D_{pq}^{ij} \leftarrow D_{pq}^{ij} + \eta \tanh\left[\beta\left(\frac{H_{pq}^{ij}}{H_{\mathrm{MM}}} - 0.5\right)\right] \tag{7}$$

$$C_{ij}^{kl} \leftarrow C_{ij}^{kl} - \tau \tanh\left[\alpha\left(\frac{H_{ij}^{kl}}{H_{\mathrm{MM}}} - 0.5\right)\right]. \tag{8}$$

In these equations, H_{MM} is the maximum possible value of the joint entropy between two CPEs. The parameters η, τ, β, and α define the rate of learning. We modify the coupling coefficients D_{pq}^{ij} and C_{ij}^{kl} based on the value of joint entropy between two processing elements. Joint entropy of two logistic maps, as it is shown in Fig. 6, is maximum when the bifurcation parameters of both maps are maximized and it is minimum when the maps are in the fixed attractor regime (lower values of the bifurcation parameters). In this regard, the joint entropy resembles the well-known Hebbian learning rule where the coupling gets stronger if both pre- and postsynaptic neurons are persistently active and vice versa.

IV. Sparse Coding and Population Kurtosis

As it is shown by Field [32], the self-organization maps of the visual system produce sparse codes by forming more efficient representations that are suited for cognition, perception, and addressing the associative memory. Sparse coding is a ubiquitous strategy employed in several different modalities across different organisms. There is evidence that sparseness constitutes a general principle of sensory coding in the nervous system and there are experimental observations of sparse coding in brain [4]. Sparse coding has benefits, namely, it increases the storage capacity in associative memory, it simplifies the process of read out in subsequent levels of processing, etc. [13], [4].

To assess the sparseness of a code that is produced by the network, one can use the population kurtosis [4], [41] or the Treves–Roll population sparseness [38]. In this paper, we use the population kurtosis to measure sparseness of representations. Population kurtosis is the kurtosis of the response of the entire population of CPEs to a single stimulus. In a network of $N \times N$ CPEs, the population kurtosis of a representation measures the fourth statistical moment relative to the variance squared

$$\mathrm{Kurt} = \left\{\frac{1}{N^2}\sum_{i,j}\left[\frac{\Omega_{ij} - \bar{\Omega}}{\sigma}\right]^4\right\} - 3 \tag{9}$$

where in our model, $\Omega_{ij} = H_{ij}/H_M$ is the normalized entropy of the activity of the ijth CPE. Other variables $\bar{\Omega}$ and σ are the mean response and standard deviation of Ω_{ij}, respectively.

V. Numerical Simulations

In this section, we review the results that are obtained by numerical simulation of the network in two separate examples. First, we stimulate the network with face images and we see how the network produces sparse representations. Then we continue by exciting the network with a set of natural images and

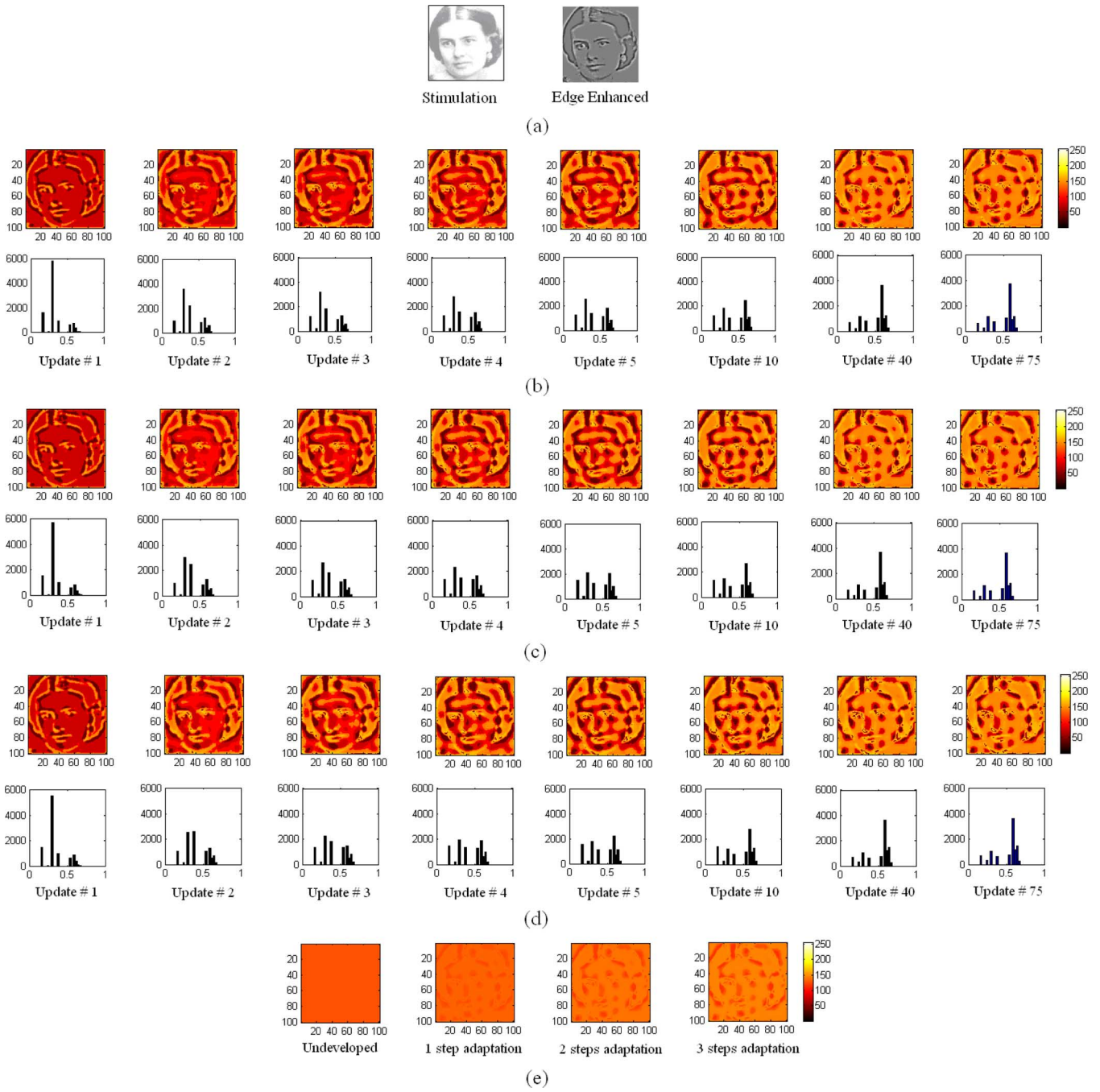

Fig. 8. Nonlinear coupling using the function $f(x, y) = x^y$. In this function, a CPE is very sensitive to a particular stimulus even when the excitation is weak (small y) or not sensitive to a stimulus unless the excitation is very strong (large y). See Section III-B for further explanation.

we show how the cortical computational maps are generated in this network.

Example 1: Stimulation of a PCLMN With Face Images

In our first example, a cortical patch is stimulated by processing elements of early processing layers. For instance, suppose that a 100×100 image, shown in Fig. 8(a)(left), is projected on the retina. The early processing layers, which are the receptive fields of the ganglion cells in retina and the neurons in LGN, behave as edge enhancer filters. The edge enhanced image, shown in Fig. 8(b)(right), stimulates an undeveloped cortical patch of 100×100 CPEs. In the undeveloped network, all the coupling factors D_{pq}^{ij}s and C_{ij}^{kl}s are initially set to 1.0 so that the couplings are linear. Parameters of the dynamics are set as follows: $\xi = 0.18, \gamma = 0.15, \lambda = 0.18, \eta = 0.1, \beta = 1.0$. In order to simplify the computations, in this example, we assume $A = 1$, so that each CPE in the processing layer is stimulated by one CPE in the stimulation layer. After stimulation, we iterate and update the network and monitor the patterns of activities of all processing elements. The evolution of the normalized entropies of the activities of processing elements of the cortical patch are given in the form of the sequence of images and histograms in Fig. 8(b). In these images, the darker parts show the areas that are strongly stimulated and the dynamics of the CPEs

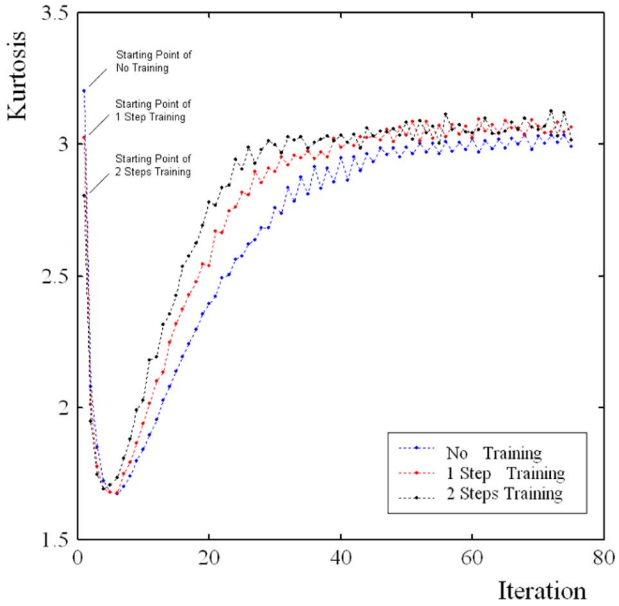

Fig. 9. Evolution of the population kurtosis as a function of the number of updates in the undeveloped network and the developed network of Example 1 after 1- and 2-steps adaptation. During first updates, the kurtosis drops rapidly, because at the beginning almost all CPEs are strongly stimulated. When initial moments are passed, the sparcification process takes place that develops sparse bubbles of ordered activities. In this stage, the kurtosis increases monotonically and ultimately merges to a constant value; however, the final value of the kurtosis is almost independent of training. Training only expedites the convergence process.

in these regions are pushed down to the ordered activity regime. As we can see in the figure, at the beginning, this stimulation causes a lot of ordered activities; however, the cooperation and competition of the CPEs with their neighbors drive many CPEs to the chaotic regime and produce very sparse bubbles of ordered activities. Following each update, we compute the population kurtosis and continue the iteration-update process until the changes in the population kurtosis become negligible. Evolution of kurtosis for this example is shown in Fig. 9. As we can see in this figure, kurtosis has a higher value at first, which occurs because at the beginning most CPEs are pushed to the ordered activity regions [first histogram in Fig. 8(b)]. Nevertheless, the first few updates reduce the sparsity of the activity distribution. Continuing the process, the sparsity of the representation and the corresponding kurtosis improves when the ordered activity bubbles are being developed. This can be seen clearly in the histograms given in Fig. 8(b).

When the change in the kurtosis is negligible, it is time for the network to self-organize itself by adapting the coupling coefficients. For simplification purposes, in this example, we only train the afferent coupling coefficients. The effect of adaptation on the afferent coupling coefficients is displayed in Fig. 8(e). As we can see in this figure, the network organizes itself based on the stimulations that it receives from the surrounding environment. After adaptation, if one stimulates this cortical network with the same input pattern, as it is displayed in Fig. 8(b), the modified coupling coefficients improve the selectivity of the network and expedite the sparsification process so that the kurtosis takes a lower value at the beginning and converges to the final sparse representation faster. The second step adaptation [Fig. 8(d)] improves the performance of the network further. Comparison of the evolution of kurtosis before and after adaptation (Fig. 9) shows that the final sparse codes that are generated by both undeveloped and developed networks are almost the same and the difference between the two networks is manifested in the speed of two systems. Another subtle point in the kurtosis diagram is that the first adaptation improves the speed of the system considerably; however, the second adaptation does not improve the speed as much as the first one did and so on. Consequently, more adaptations cannot improve the speed of the system unlimitedly and overtraining of the network is not useful. Results of this numerical simulation have been tested for different stimulations; the behavior of the network was quite similar.

To continue this example, we stimulate another processing cortical patch, with the same parameters as the first one, and with the sparse code that is developed by the first cortical processing network (Fig. 10). As we can see in Fig. 10, the sparsification process continues in the second layer and after few steps of iteration updates, the produced entropy pattern possesses very high peakedness. Now, if we apply a threshold to the entropy patterns to discern only the bubbles of the highest ordered activities, the result will be a very sparse code that is reminiscent of the theory of gnostic units or grandmother cells [40] [Fig. 10(d)]. This theory says that initial areas of the visual cortex code elementary features. Then, the outputs of these neurons combine to produce higher order selective detectors so that, at the upper levels, neurons of the inferior temporal lobe produce very selective responses to complex stimuli [39]. This is in agreement with our hierarchical model presented here. As we can see in this example, higher cortical patches become more selective and the fraction of CPEs with ordered activities compared to the fraction of CPEs with chaotic patterns of activity reduces in the higher processing layers. In addition, our simulations show that (to some extent) the developed gnostic unit/units produced by the network are invariant with respect to scaling and addition of noise.

Example 2: Stimulation of a Cortical Patch With Images of Nature

In our second example, we stimulate a cortical patch with early processing layers that preprocess a set of 512×512 pixel images taken from nature. However, preprocessing in this case includes both applying optimal whitening filter [2] and edge enhancement (Fig. 11). Same as the previous example, we start with an undeveloped network in full chaotic regime. Parameters of the dynamics are set as follows: $\xi = 0.1, \gamma = 0.2, \lambda = 0.4, \eta = \tau = 0.005, \alpha = 1.0, \beta = 3.0$. In the first part of this example, we assume that $A = 1$ but we increase it later. After stimulating the network with a preprocessed image, we apply the iteration-update process and simultaneously monitor the kurtosis until its changes become negligible. Then, we do the adaptation by modifying the afferent coupling coefficients D_{pq}^{ij}. After adaptation, we take another image and follow the same protocol and we repeat this process for 50 different images. After 50 steps of adaptation, the afferent coupling matrix will look like the image displayed in Fig. 12(a). Now, if we take this afferent coupling matrix and compute the sensitivity of each

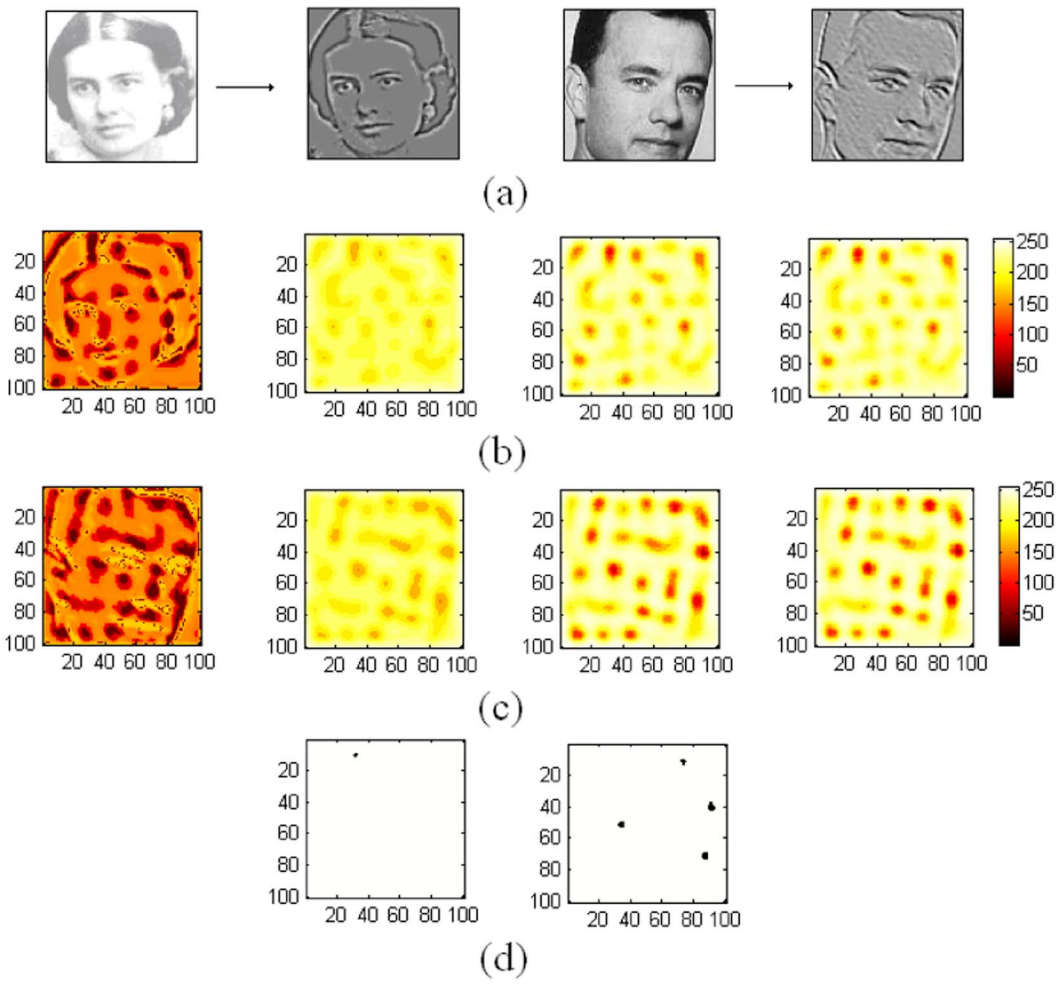

Fig. 10. Stimulation of a second cortical patch by the sparse representations that are generated in the first cortical processing layer. The second processing layer produces highly sparse codes so that if we apply a threshold level to extract only the bubbles of the highest ordered activities, the output will be the gnostic unit/units. (a) Input images and the edge enhanced patterns. (b) and (c) The left most pattern is the sparse input that feeds the second processing layer. Others show the evolution of the normalized entropies of the activities of the CPEs in the second processing layer after 10, 20, and 30 iterations. (d) Gnostic unit/units for the woman's image (left) and the man's image (right). Parameters of the second processing layer are the same as the first layer given in the caption of Fig. 8.

element of this matrix to different orientations, what comes up is given in Fig. 12(b). As we can see in this figure, elements with the same orientation preferences cluster together and produce the iso-orientation patches, which are similar to the ones observed experimentally. If we magnify a patch from this orientation preference map, we can discern the linear zones, pin wheels, fractures, and saddle points that occasionally appear in the orientation preference maps that are obtained by other experimental observations or theoretical modelings [6].

To investigate details of the adaptation process, we study a case where the cardinality of the set of afferent inputs to each processing CPE A is more than one. For instance, we assume that an early processing layer, which includes whitening and edge enhancement of natural images, stimulates a processing layer. We start with an undeveloped network where $A = 17 \times 17, E = 49, I = 240$, and the size of the network is 100×100 CPEs. After stimulation, we carry out the iteration-update process repeatedly until convergence; then we follow the adaptation rule for the afferent matrix given in (7) and we perform this for 1000 natural images. The adaptation process, step by step, modifies the coupling coefficients so that massively trained coefficients change from uniform state without any specific preference to some localized, bandpass, orientation preference wavelets. Samples of such coupling coefficients for every tenth CPE from the array of 100×100 units are displayed in Fig. 13(a). Few sample receptive fields with similar orientation preference but different frequency selectivity are shown in Fig. 13(b). Fig. 13(c) shows few sample receptive fields with different orientation preferences. One can use the conventional drifting grating method to measure the preferred spatial frequency and orientation of the receptive field of each CPE in the network [43].

VI. Conclusion

In this paper, a new model for the way cortical patches process information and interact with subcortical areas was presented and the capabilities of such networks in producing sparse codes and computational maps were studied. In contrast to many previously presented models where the processing elements are simple, such as sigmoidal neurons, processing elements in this work were nonlinear quadratic maps (logistic maps) that have

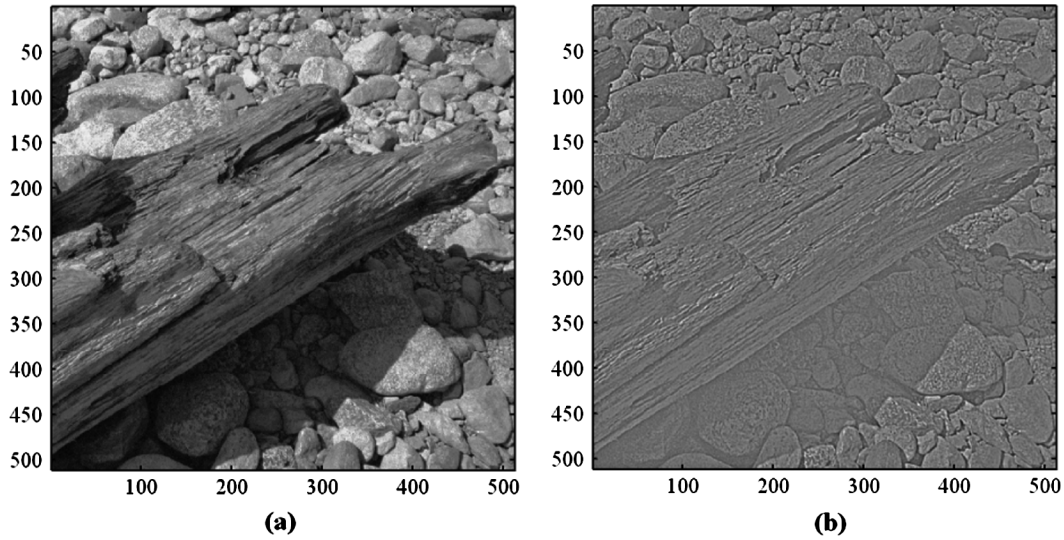

Fig. 11. (a) Typical image from nature that is used in Example 2. (b) The preprocessed image obtained by applying the optimal whitening and edge enhancing filters to the original image from nature. Optimal whitening filter: Atick [33] and van Hateren [34], [35] formulated the theory of coding in retina based on whitening the power spectrum of natural images in space and time and Field [1] has shown that natural scenes posses a characteristic $1/f^2$ spatial power spectrum. Consequently, the optimal whitening filter for natural images has a transfer function that rises linearly with spatial frequency and falls off where the signal power becomes equal to or less than the noise power. In this example (same as [2]), we take a raw image, then we apply the zero-phase whitening/low-pass filter $R(f) = f \times \exp(-(f/f_0)^4)$ where $f_0 = 200$ cycles/picture. Part of the natural images used in our work (including the one that is shown in this figure) are adapted from a database kindly provided by B. Olshausen and can be found as part of the *sparsenet* software package [42].

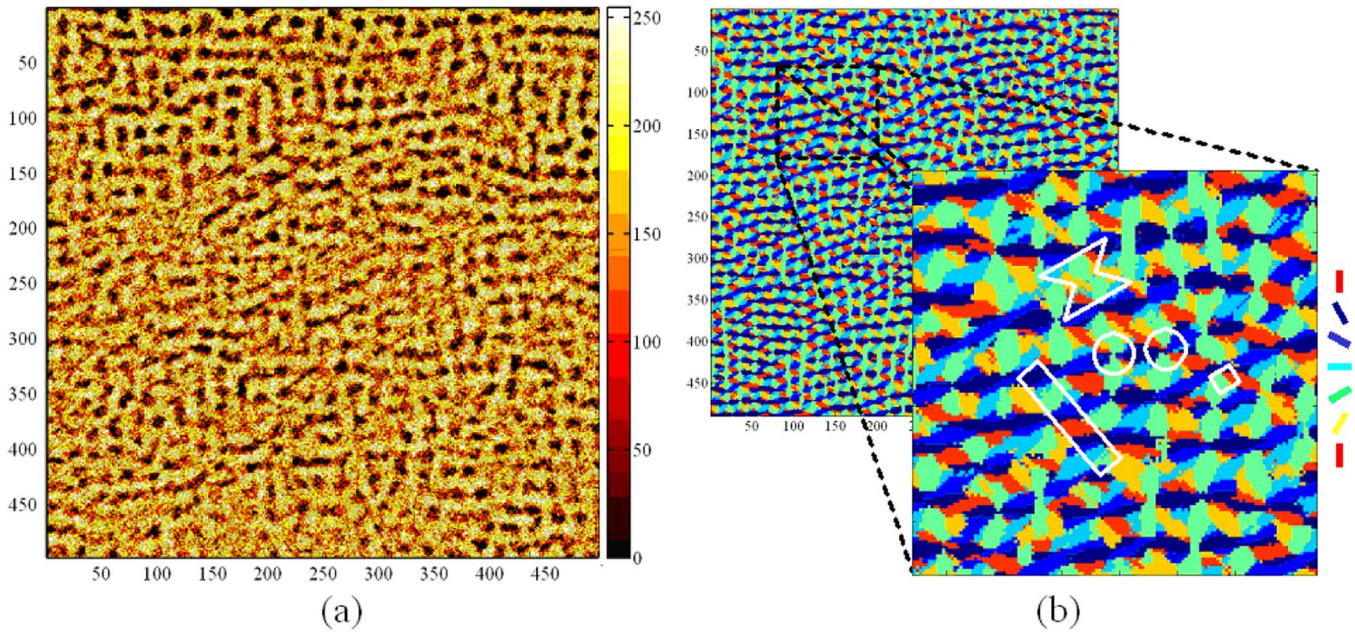

Fig. 12. (a) Massively trained D matrix. (b) Orientation preference map produced by the PCLMN. If we magnify part of this map we can see that nearby CPEs prefer the same orientation forming groups known as iso-orientation patches. A linear zone (continuous change of orientation along a straight line, rectangle), a pair of pinwheels (where orientation changes continuously around a point, circles), a saddle point (long patch of one orientation, bowtie), and a fracture (sharp transitions from one orientation to a very different one, square) are highlighted in the magnified plot. The notations that are used in this example to mark various zones are adapted from [6]. The process starts with $D_{ij} = C_{ij} = 1.0$ for all is and js and: $\xi = 0.1, \gamma = 0.2, \lambda = 0.4, \eta = \tau = 0.005, \alpha = 1.0, \beta = 3.0$ and $A = 1, E = 49, I = 240$.

rich dynamics that includes fixed-point attractors as well as periodic and chaotic behavior. This model for processing elements was derived by applying some modification to the netlet hypothesis previously reported by Anninos and Harth. In our next step, considering the experimental results developed in Freeman's laboratory, processing elements were coupled parametrically to form PCLMN. A set of equations for training and adaptation of the network was given as well. Simulation of such a network and coupled networks of this kind show unique capabilities of the PCLMN in producing sparse codes and cortical computational maps such as orientation preference maps as well as capabilities for producing localized, frequency selective, and orientation preference receptive fields that are reminiscent of the well-known Gabor filters. All these results were developed by exploiting real-world stimulations such as face images and a database of natural images. Despite the fact that direct compar-

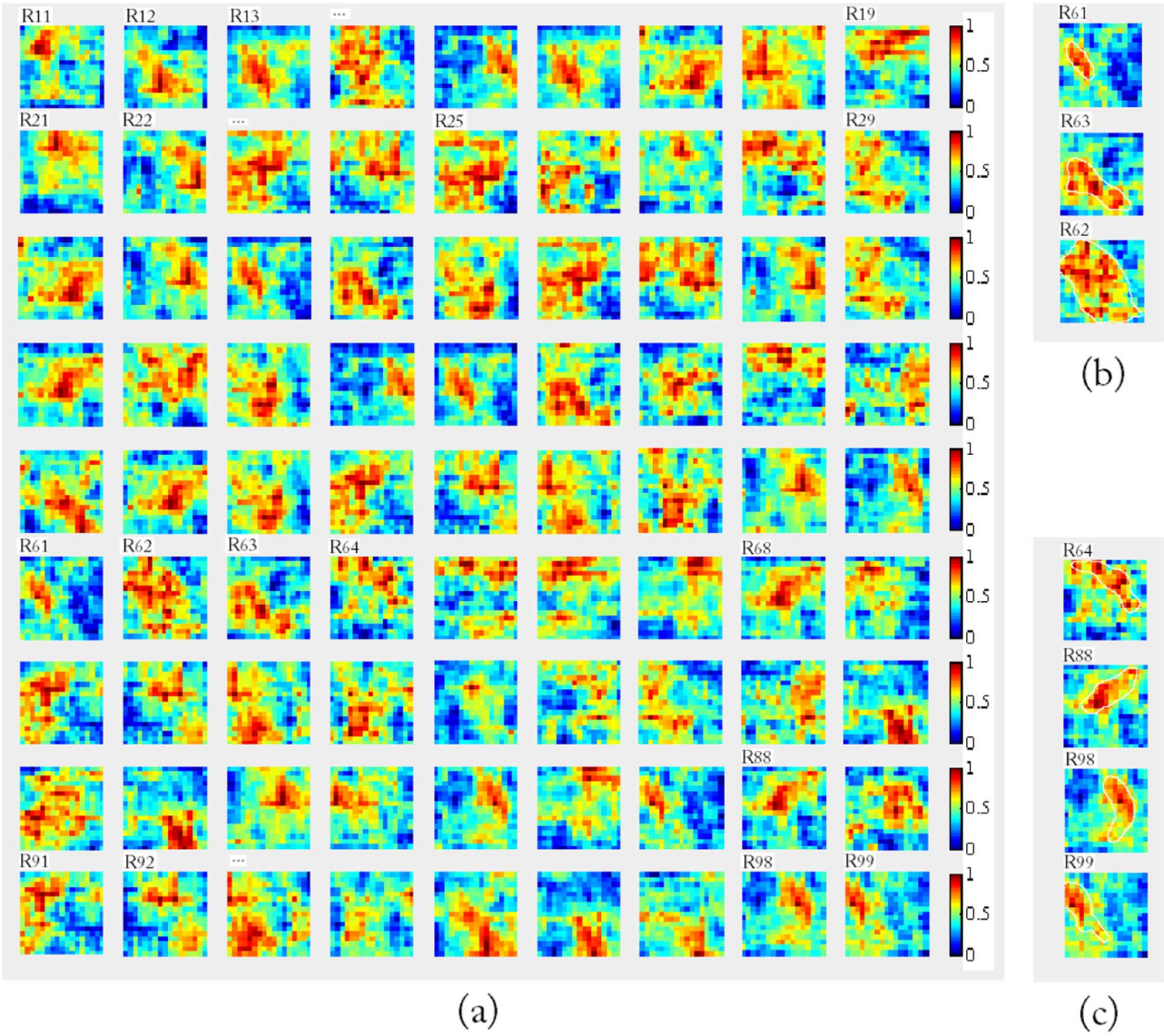

Fig. 13. (a) Typical set of receptive fields or wavelets that are produces by massive adaptation of the afferent coupling coefficients. The number of CPEs in the referenced network is 100×100 and sample 17×17 receptive fields of every tenth CPE is shown in the figure. As one can see in this figure, the many of these receptive fields are localized, bandpass, with specific orientation preference. Large arrays of such receptive fields can span the entire image space. (b) Few sample receptive fields with similar orientation preference but sensitive to different spacial frequencies. (c) Few sample receptive fields with different orientation preferences.

ison of this model with other models is not practical, our simulations show that the new model, as a result of the high nonlinear functionalities involved, produces the sparse codes and computational maps considerably faster than other models reported previously. (The simulations reported in this paper were performed on a personal computer with 1 GB of random access memory and the runtime was in the order of 1–2 h.) This paper may also be used for further study of a model of this kind to produce other cortical maps (e.g., ocular dominance maps) and to find reasonable justification for certain brain activities.

REFERENCES

[1] D. J. Feild, "Relations between the statistics of natural images and the response properties of cortical cells," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 4, pp. 2379–2394, 1987.

[2] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 13, pp. 607–609, 1996.

[3] B. A. Olshausen, "Principles of image representation in visual cortex," in *The Visual Neuroscience*, L. M. Chalupa and J. S. Werner, Eds. Cambridge, MA: MIT Press, 2003, pp. 1603–1615.

[4] B. A. Olshausen and D. J. Field, "Sparse coding of sensory inputs," *Current Opinion Neurobiol.*, vol. 14, pp. 481–487, 2004.

[5] J. A. Bednar and R. Miikkulainen, "Learning innate face preferences," *Neural Comput.*, vol. 15, no. 7, pp. 1525–1557, 2003.

[6] R. Miikkulainen, J. A. Bednar, Y. Choe, and J. Sirosh, *Computational Maps in the Visual Cortex*. New York: Springer-Verlag, 2005.

[7] J. A. Bednar and R. Miikkulainen, "Tilt after effects in a self-organizing model of the primary visual cortex," *Neural Comput.*, vol. 12, pp. 1721–1740, 2000.

[8] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biol. Cybern.*, vol. 43, pp. 59–69, 1982.

[9] N. H. Farhat, "Corticonics: The way to designing machines with brain-like intelligence," *Proc. SPIE, Critical Technol. Future Comput.*, vol. 4109, pp. 103–109, 2000.

[10] N. H. Farhat, "Coticonic networks for higher-level processing," in *Proc. 2nd Int. Assoc. Sci. Technol. Develop.*, Feb. 23–25, 2004, pp. 256–262.

[11] N. H. Farhat, "Corticonic models of brain mechanisms underlying cognition and intelligence," *Phys. Life Rev.*, vol. 4, pp. 223–252, 2007.

[12] F. Alexandre, F. Guyot, J. Haton, and Y. Burnod, "The cortical column: A new processing unit for multilayered networks," *Neural Netw.*, vol. 4, pp. 15–25, 1991.
[13] P. Kanerva, *Sparse Distributed Memory*. Cambridge, MA: Bradford Books of MIT Press, 1988, ISBN-13: 978-0262111324.
[14] E. M. Harth, T. J. Csermely, B. Beek, and R. D. Lindsay, "Brain function and neural dynamics," *J. Theor. Biol.*, vol. 26, pp. 93–120, 1970.
[15] P. Anninos, B. Beek, E. Harth, and G. Pertile, "Dynamics of neural structures," *J. Theor. Biol.*, vol. 26, pp. 121–148, 1970.
[16] E. Harth, "Order and chaos in neural systems: An approach to the dynamics of higher brain functions," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-13, no. 5, pp. 782–789, Sep./Oct. 1983.
[17] R. Pashaie and N. H. Farhat, "Dynamics of electron-trapping materials under blue light and near-infrared exposure: An improved model," *J. Opt. Soc. Amer. B, Opt. Phys.*, vol. 24, no. 7, pp. 1942–1941, Aug. 2007.
[18] R. Pashaie and N. H. Farhat, "Analytic model for the dynamics of electron trapping materials with applications in nonlinear optical signal processing," *J. Opt. Soc. Amer. B, Opt. Phys.*, vol. 25, no. 1, pp. 15–21, Jan. 2008.
[19] E. Hernandez, G. Lee, and N. H. Farhat, "Analog realization of arbitrary one-dimensional maps," *IEEE Trans. Circuits Sys.*, vol. 50, no. 12, pp. 1538–1547, Dec. 2003.
[20] N. F. Rulkov, I. Timofeev, and M. Bazhenov, "Oscillations in large-scale cortical networks: Map-based model," *J. Comput. Neurosci.*, vol. 17, pp. 203–223, 2004.
[21] W. Freeman and C. A. Skarda, "Spatial EEG patterns, nonlinear dynamics and perception: Neo-Sherringtonian view," *Brain Res. Rev.*, vol. 10, p. 147, 1985.
[22] C. A. Skarda and W. J. Freeman, "How brains make chaos in order to make sense of the world," *Behav. Brain Res.*, vol. 10, pp. 161–195, 1987.
[23] P. Faure and H. Korn, "Is there chaos in brain? I. Concepts of nonlinear dynamics and methods of investigation," *C. R. Acad. Sci. Paris, Ser. III 324*, pp. 773–793, 2001.
[24] H. Korn and P. Faure, "Is there chaos in brain? II. Experimental evidence and related models," *Comptes Rendus Biol.*, vol. 326, pp. 787–840, 2003.
[25] G. Lee and N. H. Farhat, "The bifurcating neuron network 1," *Neural Netw.*, vol. 14, pp. 115–131, 2001.
[26] G. Lee and N. H. Farhat, "The bifurcating neuron network 2: An analog associative memory," *Neural Netw.*, vol. 15, pp. 69–84, 2002.
[27] G. Lee and N. H. Farhat, "Parametrically coupled sine map networks," *Int. J. Bifurcat. Chaos*, vol. 11, no. 7, pp. 1815–1834, 2001.
[28] G. Tanaka and K. Aihara, "Multistate associative memory with parametrically coupled map networks," *Int. J. Bifurcat. Chaos*, vol. 15, no. 4, pp. 1395–1410, 2005.
[29] N. H. Farhat, "Biomorphic dynamical networks for cognition and control," *J. Intell. Robot. Syst.*, vol. 21, pp. 167–177, 1998.
[30] K. Kaneko, "Overview of the coupled map lattices," *Chaos*, vol. 2, pp. 279–282, 1992.
[31] K. Kaneko, *Theory and Application of Coupled Map Lattices*. New York: Wiley, 1993, pp. 1–49, ISBN-13 978-0471937418.
[32] D. J. Field, "What is the goal of sensory coding?," *J. Neural Comput.*, vol. 6, pp. 559–601, 1994.
[33] J. J. Atick and A. N. Redlich, "What does the retina know about natural scene?," *Neural Comput.*, vol. 4, pp. 196–210, 1992.
[34] V. Harteren, "A theory of maximizing sensory information," *Biol. Cybern.*, vol. 68, pp. 23–29, 1992.
[35] V. Harteren, "Spatiotemporal contrast sensitivity of early vision," *Vis. Res.*, vol. 33, pp. 257–267, 1993.
[36] K. P. Balanda and H. L. MacGillivray, "Kurtosis: A critical review," *Amer. Statistician*, vol. 42, no. 2, pp. 111–119, 1988.
[37] E. T. Rolls and M. J. Tovee, "Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex," *J. Neurophysiol.*, vol. 73, no. 2, pp. 713–726, 1995.
[38] E. T. Rolls and M. J. Tovee, "Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex," *J. Neurophysiol.*, vol. 73, pp. 713–726, 1995.
[39] M. S. Gazzaniga, R. B. Ivry, and G. R. Mangun, *Cognitive Neuroscience, the Biology of the Mind*, 2nd ed. New York: Norton, pp. 210–.
[40] C. G. Gross, "Genealogy of the grandmother cell," *Neuroscientist*, vol. 8, no. 5, pp. 512–518, 2002.
[41] B. Willmore and D. J. Tolhurst, "Characterizing the sparseness of neural codes," *Comput. Neural Syst.*, vol. 12, pp. 255–270, 2001.
[42] B. A. Olshausen, The Sparsenet Software Package for MATLAB [Online]. Available: http://redwood.berkeley.edu/bruno/sparsenet/
[43] D. L. Ringach, M. J. Hawken, and R. Shapley, "Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences," *J. Vis.*, vol. 2, no. 1, pp. 12–24, Jan. 2002.

Ramin Pashaie (S'07) was born in Tehran, Iran. He received the B.S. degree in electrical engineering (with a major in microelectronics) from Shahid Beheshti University (SBU), Tehran, Iran, in 1998, the M.S. degree in fields and waves from Khaje Nasir Tousi University of Technology (KNTU), Tehran, Iran, and the Ph.D. degree in electrical and system engineering from University of Pennsylvania, Philadelphia, in 2007.

Currently, he is continuing his studies as a Postdoctoral Scholar at the Bioengineering Department, Stanford University, Stanford, CA.

Nabil H. Farhat (S'58–M'63–SM'72–F'81–LF'99) received the B.Sc. degree from the Technion, Haifa, Israel, in 1957, the M.Sc. degree from the University of Tennessee, Knoxville, in 1959, and the Ph.D. degree from the University of Pennsylvania, Philadelphia, in 1963, all in electrical engineering.

In 1964, he joined the Faculty of the Moore School of Electrical Engineering, University of Pennsylvania, where he is now Professor of Electrical and Systems Engineering and heads the Electro-Optics and Photonic Neuroengineering Laboratory. His current research interests are in collective nonlinear dynamical information processing, neural networks, photonic realization of neurocomputers, and corticonics where he is applying concepts and tool from nonlinear dynamics, bifurcation theory, information driven self-organization, and chaos to the modeling and study of cortical information processing. His teaching includes courses in neurodynamics and neural networks, electromagnetics (EM) theory, electro-optics, electron and light optics, and holography on both graduate and undergraduate levels. His past research included microwave diversity imaging, holography, automated target recognition, optical information processing, and the study of the interaction of EM radiation with plasma and solids in the context of millimeter wave and laser output energy measurement with glow discharge plasma.

Dr. Farhat has held the Ennis Chair in Electrical Engineering, served as a Distinguished Visiting Scientist at the Jet Propulsion Laboratory, Pasadena, CA, is a recipient of the University of Pennsylvania Christian R. and Mary F. Lindback Foundation Award for Distinguished Teaching and is a Fellow of the Optical Society of America, member of the Electromagnetics Academy, the American Institute of Physics, Sigma Xi, and Eta Kappa Nu. He has also served on the National Board of Directors of Eta Kappa Nu, as an RCA Consultant, and as an Editor of *Advances in Holography*, Associate Editor of *Acoustical Imaging and Holography*, Action Editor of *Neural Networks* and *Neural Computation* and Advisory Editor of *Optics Letters*.