Pattern Association and Retrieval in a Continuous Neural System

Hung-Jen Chang and Joydeep Ghosh [1]

Department of Electrical and Computer Engineering
The University of Texas at Austin
Austin, TX 78712

June 18, 1992

Abstract—*This paper studies the behavior of a large body of neurons in the continuum limit. A mathematical characterization of such systems is obtained by approximating the inverse input-output nonlinearity of a cell (or an assembly of cells) by three adjustable linearized sections. The associative spatio-temporal patterns for storage in the neural system are obtained by using approaches analogous to solving space-time field equations in physics. A noise-reducing equation is also derived from this neural model. In addition, conditions that make a noisy pattern retrievable are identified. Based on these analyses, a visual cortex model is proposed and an exact characterization of the patterns that are storable in this cortex is obtained. Furthermore, we show that this model achieves pattern association that is invariant to scaling, translation, rotation and mirror-reflection.*

[1]Please contact *hjchang@pine.ece.utexas.edu* or *ghosh@pine.ece.utexas.edu* for any questions or comments regarding this paper, or for reprint requests.

1 Introduction

Biological neural networks consist of very large numbers of neurons. The human brain has over 10^{11} neurons, with an average connectivity in the thousands. Faced with numbers of this magnitude, it is impossible and often meaningless to monitor every single neuron as it interacts with its surroundings. To gain insight into the complex functions performed by neural systems, a description of the overall network behavior is required that is not overwhelmed by detailed behavior of individual neurons. This problem can be tackled by a two-step approach. The first step is to deal with the interaction of numerous individual cells in one homogeneous neuron-assembly using statistical methods [Ama72a, Ede87, CGL91]. The second step is to deal with the interaction of numerous homogeneous assemblies by way of a neural field or continuum theory [Ama77].

In this paper, we focus on the second step by considering a high-density neural network as a continuous "field", wherein the discrete subindex used to label a particular neuron (or assembly of neurons) is replaced with a continuous position vector. Although the penalty is that we need an additional variable to identify the neural quantities, the system description gets reduced from numerous coupled differential equations to only one intergo-partial-differential equation. By solving the derived spatio-temporal equation, we can get a good picture of the neural system.

The continuous neural model studied in this paper is based on a somewhat detailed characterization of individual neurons stemming from an associative network model with a time delay factor. This model is used to deal with the problems of storable pattern formation and its retrieval in a high-density neural system. The primary concerns of this investigation are to determine the patterns associated with the neural system and the conditions required for such patterns to be retrievable from some noisy versions. Under the hypothesis of *adiabatic learning* [Cai61], expressions for key quantities, such as output voltage are obtained. The results are applied to analyze a visual cortex model with "on-center, off-surround" interconnections and sigmoidal input-output transfer function. An exact solution of the associative storable output activity is achieved and is used to explain several known characteristics of a visual system, such as visual invariances to pattern scaling, translation, rotation and

mirror-reflection.

Section 2 summarizes related research on neural continuum models and their applications to pattern generation, pattern retrieval and visual invariances. In the next section, starting from a discrete neural model of associative memory, a continuous model of neural networks is presented and its corresponding pattern generation equation is obtained. Then, the noise-reducing equation is derived. These mathematical results are used in Section 4 to analyze a visual cortical system with Mexican-hat type interconnections and sigmoidal input-output transfer function. The characteristic of invariances with respect to an affine transformation is presented in Section 5.

2 Background

2.1 Neural Continuum Models

There have been several neural field studies (see, e.g., Wiener and Rosenblueth, 1946 [WR46]; Beurle, 1956 [Beu56]; Griffith, 1963, 1965 [Gri63, Gri65]; Farley and Clark, 1961 [FC61]; Coleman, 1971 [Col71]; Wilson and Cowan 1973 [WC73]; Boylls 1975 [Boy75]; Ellias and Grossberg, 1975 [EG75]; Stanley, 1976 [Sta76]; Amari, 1977 [Ama77]; Amari and Arbib 1977 [AA77]; Ermentrout and Cowan 1980 [EC80]). All these researchers considered a 1-d or 2-d cortical domain. Then, they argued that the interconnection strengths are dependent only on the distance between connections and additionally assumed that the neural tissue is isotropic. Under these conditions, they took the continuum limit of a discrete associative neural model [Hop84] (see also Eq. (1)), assuming there are infinitely many neurons distributed on a cortical surface. In this manner, they obtained a version of a continuous neural field of the form:

$$C\frac{du(x,y,t)}{dt} = -\frac{1}{R}u(x,y,t) + \int_{\mathcal{X}}\int_{\mathcal{Y}} w(x,y)v(x,y,t)dx\,dy + I^{(s)}(x,y,t).$$

By way of the similar technique of neural field, Wilson and Cowan [WC73] proposed a plausible field equation and showed by computer simulations that their model has three interesting types of pattern dynamics. Reverberation and propagation of excitation patterns was studied in discrete homogeneous nerve networks [Ama75]. Boylls [Boy75] proposed an

interesting field model of the cerebellum. Amari also studied the neural layer characteristics of certain types of pattern dynamics in a 1-d case [Ama77], and subsequently, extended the earlier model to a 2-d competitive-cooperative neural field [AA77]. Using some simplifying assumptions such as ignoring the time lag, Amari *et al.* [Ama77, AA77] studied the quantity of average membrane potential (input voltage) in a macroscopic sense. They gave an analysis of the equilibrium states of the system and their stability, and the interaction of different stimuli separated in space and time. These theoretical foundations are explored by graph analysis in phase space. However, there is no explicit mathematical expression for the associative spatio-temporal patterns, and also no mechanism is proposed for their retrieval.

2.2 Pattern Generation

The concept of a neural field was applied to biological pattern generation by Ermentrout and Cowan [EC79], who considered an excitatory and inhibitory neural system and studied a time coarse-grained current or firing rate. By making a linear approximation to a sigmoidal type transfer function, they showed the existence of a variety of doubly-periodic patterns, such as hexagons, rolls, etc., as solutions to 2-d neural networks under the continuum limit.

In 1980, a generalized model for the spatio-temporal activity of neural networks was proposed [EC80]. In this work, stationary periodic spatial patterns are discussed from the view-point of bifurcation theory. Existence of spatial patterns on the whole line is established by the implicit function theorem. Singularity theory is used to study the local structure of the bifurcation equations. A Poincaré-Lindstedt series is developed to establish the form of the periodic stationary states and their stability. The biological relevance of these patterns is also discussed but the biological significance of these spatially periodic solutions is still not clear. Some of these results are obtainable as a special case of the neural field model developed in this paper.

Ermentrout and Campbell [Erm86] also examined an excitatory and inhibitory neural system and considered both discrete time and continuous time models. They studied the quantity of pigment secretion and assumed that it is stimulated by neural activity or is simply proportional to neural activity. Then, they presented lots of spatial patterns on the shells of molluscs by computer simulations. Nevertheless, they did not give a general

close form expression for the possible spatio-temporal patterns, and again, no mechanism is proposed for the retrieval.

2.3 Pattern Retrieval and Visual Invariances

The use of artificial neural networks as distributed associative memories, and the associated problem of retrieval of memorized patterns from noisy inputs, has a long history [Ama72b, And72, Koh72, HA81, Hop82]. For instance, Anderson used various kinds of filters designed so as to minimize the probability of error by maximizing a certain signal-to-noise ratio [And68, And70], and Kohonen used the concept of adaptive filter operation for autoassociative and heteroassociative recall [Koh89].

Another related work is invariant pattern recognition. In general, we can classify techniques for invariant pattern recognition into three categories: The first one is by way of invariant feature spaces. It has long been recognized that it is possible to extract features (such as Zernicke moments) which are invariant to transformations (e.g. rotations) of an input [Kan88]. The next category is invariance by training, wherein a number of different examples of the same object are directly presented to the neural networks during training. This approach was taken by Rumelhart, Hinton, and Williams [RHW86] to obtain rotational invariance for simple 3×3 templates which they investigated, but is considered impractical for real life patterns. The final category is invariance by structure. The idea of imposing invariance on a neural network by placing limitations on its structure has been proposed by various authors [Fuk80, WHHL89].

Roughly speaking, our scheme belongs to the third category. In this paper, we investigate the application of a continuous neural system to invariant pattern retrieval from a similar consideration but using a novel approach.

3 Mathematical Analysis of Spatio-Temporal Neural Systems

In this section we develop and analyze a continuous neural field model which can associate spatio-temporal patterns.

A discrete neural model shown in Eq. (1) has been widely used to describe the dynamics of n interconnected networks of neural-like elements [KS88, Hop84]:

$$C\frac{du_i(t)}{dt} = -\frac{1}{R}u_i(t) + b\sum_{k=1}^{n} w_{ik}(t)\left[v_k(t)+1\right] + I_i^{(s)}(t). \tag{1}$$

In Eq. (1), $u_i(t)$ denotes the membrane voltage, of the ith neuron or cell at time t, and $v_i(t)$ is its corresponding output voltage obtained through a sigmoidal transfer function with range from -1 to 1. A dimensionless constant b is introduced to rescale the range of output voltage for comparisons with biological models. In the absence of input to this cell, its output voltage decays exponentially to 0 (the arbitrarily chosen "resting-voltage") with a passive time constant $\frac{1}{a} = RC$, due to the cell membrane behaving as an resistive-capacitative circuit. The "conductance" $w_{ik}(t)$ of synapse from k to i, may be positive (excitatory) or negative (inhibitory). $I_i^{(s)}(t)$ represents the external input stimulus (current). Recollect that the system variables and parameters in Eq. (1) can alternatively be interpreted as the average values in a homogeneous cell assembly from the neural ensemble theory [Ama72b, CGL91]. In that case, Eq. (1) describes the behavior of interactions among n neural assemblies.

We are concerned here with seeing how and what kinds of associative output voltage patterns are generated in response to the external input stimulus. First, we introduce the inverse input-output transfer function $f\left(v_i(t+\Delta t)\right) = u_i(t) - \theta(t)$, taking into account for the time delay factor Δt occurring in the transfer of input-output activity, and the threshold $\theta(t)$ for neuron firing. Thus, Eq. (1) becomes

$$\frac{df\left(v_i(t+\Delta t)\right)}{dt} = -af\left(v_i(t+\Delta t)\right) - a\theta(t) - \frac{d\theta(t)}{dt} + \frac{b}{C}\sum_{k=1}^{n} w_{ik}(t)\left[v_k(t)+1\right] + \frac{1}{C}I_i^{(s)}(t). \tag{2}$$

Eq. (2) is such a highly complex nonlinear equation set that there is no general method to solve it. Fortunately, by exploiting the special features of the "neuron" (or "neural assembly") itself and its "connectivity," we can still solve the problem and obtain an understanding about the system's spatio-temporal behavior, as follows.

The inverse sigmoidal function is approximated by three lines, i.e., a linear middle region bounded by positive and negative saturations on either side as shown in Fig. (1). Then, we solve for cell (or cell assembly) dynamical behavior in every region separately. The

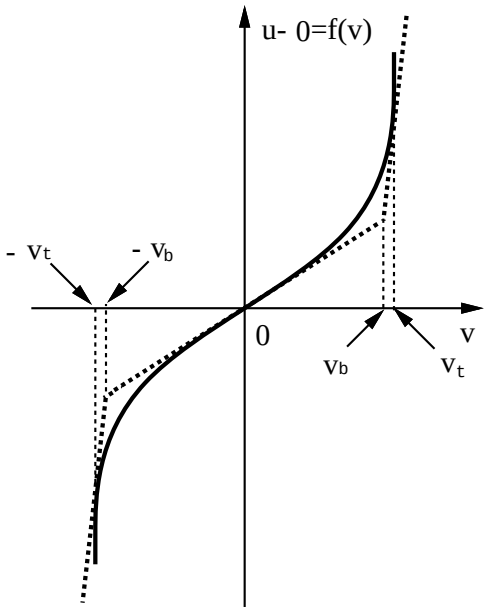

Figure 1: inverse sigmoidal input-output transfer function and its linearized sections

complete solution is obtained through combining these sub-solutions by matching boundary conditions imposed from continuity. Let us linearize the desired section of the inverse function in such a way that

$$f\left(v_i(t)\right) \approx f\left(v_t\right) + f'\left(v_t\right)\left[v_i(t) - v_t\right], \tag{3}$$

where v_t represents the x coordinate of a tangent point on curve $y = f(x)$. Eq. (2) then becomes

$$\begin{aligned} \frac{dv_i(t+\Delta t)}{dt} &= -a\,v_i(t+\Delta t) + \frac{b}{C f'\left(v_t\right)} \sum_{k=1}^{n} w_{ik}(t) v_k(t) + \frac{I_i^{(s)}(t)}{C f'\left(v_t\right)} \\ &+ \frac{1}{f'\left(v_t\right)} \left[\frac{b}{C} \sum_{k=1}^{n} w_{ik}(t) - a\theta(t) - \frac{d\theta(t)}{dt} + a f'\left(v_t\right) v_t - a f\left(v_t\right) \right]. \end{aligned} \tag{4}$$

For example, if we are interested in the middle section, we may set $v_t = 0$ so that $f\left(v_t\right) = 0$. Then the solution $v_i(t)$ of Eq. (4) corresponding to this value of v_t should lie in $(-v_b, v_b)$

(see Fig. 1). If $v_i(t)$ goes out of this range at some later time, the solution obtained is no longer valid but it indicates the region into which the system variable $v_i(t)$ has now actually entered. So we simply evaluate Eq. (4) using the new v_t. In this way, we can get the the right solution of $v_i(t)$ by using the proper linearized section in Eq. (3).

Meaningful ranges for various parameters can be obtained from biological data. We know that the brain has numerous neurons with high cell density and connectivity. Experimental data have shown that the total cortical thickness of the tissue in area 17 in the cat is on the order of 48,000 to 50,000 neurons per mm^3 [BC83], and its primary visual cortex amounts to an average of 6,000 synaptic connections per cell [BC85]. This microscopic consideration motivates a continuous version of the neural model in which the discrete cell subindex i is replaced by a continuous position vector $\vec{r}$. Thus Eq. (4) gets transformed into:

$$\begin{aligned}\frac{dv(\vec{r},t+\Delta t)}{dt} &= -a\,v(\vec{r},t+\Delta t)+\frac{b\varrho}{C\,f'(v_t)}\int_{\mathcal{D}} w(\vec{r},\vec{r}\,',t)v(\vec{r}\,',t)d\vec{r}\,'+\frac{I^{(s)}(\vec{r},t)}{C\,f'(v_t)} \\ &+ \frac{1}{f'(v_t)}\left[\frac{b\varrho}{C}\int_{\mathcal{D}} w(\vec{r},\vec{r}\,',t)d\vec{r}\,'-a\theta(t)-\frac{d\theta(t)}{dt}+a\,f'(v_t)\,v_t-a\,f(v_t)\right], \quad (5)\end{aligned}$$

where ϱ is the cell density in the whole system domain $\mathcal{D}$. We assume that this density is constant for simplification.

If a small portion of a nerve system consists of many randomly connected neurons, it can be regarded as a homogeneous neural assembly. Then, a statistical manipulation in every assembly yields a continuous field equation for the entire neural system [Ama77], and we can use a continuous position vector $\vec{r}$ to represent the location of the ith assembly of cells. This induces a neural continuum from a macroscopic view-point. By taking $\varrho \sim 1$, Eq. (5) is qualified to represent the dynamics among n neural assemblies.

Physiological observations justify the use of an *adiabatic learning hypothesis* to simplify the problem by holding connection strengths and thresholds as constants (not functions of time) when finding change in cell outputs [Cai61]. Moreover, because of short range interaction between neurons, we can also claim $\int_{\mathcal{D}} w(\vec{r},\vec{r}\,')d\vec{r}\,'$ is constant (not a function of position) if we neglect the boundary effects. This will become more clear when we deal with the visual cortex model in the next section. Throughout the discussion, the last term in Eq. (5) is approximated by a time and position independent value.

Thus, from the above facts of biological neural systems, a simplified neural model is induced and we may start to study our two main concerns, i.e., pattern generation and pattern retrieval, by mathematical techniques.

3.1 Stored Patterns

During a storing process, we interpret the entire external stimuli $I^{(s)}(\vec{r},t)$ as an input pattern to be stored, and the corresponding output activities after the time delay Δt (see Eq. (5)) as the pattern that is actually presented to the memory. These two spatio-temporal patterns could only be different by a scaling factor ρ with dimension *conductance*. Thus, we call the linearized integro-partial-differential equation, Eq. (5) a pattern-storing equation when $I^{(s)}(\vec{r},t) = \rho v(\vec{r},t+\Delta t)$. The associated homogeneous differential equation of Eq. (5) may be solved by the method of separation of variables. Let us try the form of "a product term" as follows:

$$\left.\begin{array}{rcl} v(\vec{r},t) & = & \Phi(t)\Psi(\vec{r}) \\ I^{(s)}(\vec{r},t) & = & \rho\Phi(t+\Delta t)\Psi(\vec{r}) \end{array}\right\}. \tag{6}$$

Inserting Eq. (6) into the associated homogeneous differential equation of Eq. (5) and dividing the latter by $\Phi(t)\Psi(\vec{r})$ on both sides, a pair of single-variable equations are derived:

$$\left.\begin{array}{l} \frac{1}{\Phi(t)}\frac{d\Phi(t+\Delta t)}{dt} + \left[a - \frac{\rho}{Cf'(v_t)}\right]\frac{\Phi(t+\Delta t)}{\Phi(t)} = K \\ \frac{b\varrho}{Cf'(v_t)}\int_{\mathcal{D}} w(\vec{r},\vec{r}\,')\Psi(\vec{r}\,')d\vec{r}\,' = K\Psi(\vec{r}) \end{array}\right\}, \tag{7}$$

with K a constant. A paticular solution of the non-homogeneous differential equation is given by

$$v(\vec{r},t) = \frac{b\varrho R\int_{\mathcal{D}} w(\vec{r},\vec{r}\,')d\vec{r}\,' - \theta + f'(v_t)v_t - f(v_t)}{f'(v_t) - b\varrho R\int_{\mathcal{D}} w(\vec{r},\vec{r}\,')d\vec{r}\,' - R\rho}. \tag{8}$$

From the first of Eq. set (7), the solution of the temporal part can be obtained directly by integration:

$$\Phi(t) = \exp\left\{\left[\frac{\rho}{Cf'(v_t)} - a\right](t-t_s)\right\}\Phi(t_s) + K\int_{t_s}^{t}\exp\left\{\left[\frac{\rho}{Cf'(v_t)} - a\right](t-\tau)\right\}\Phi(\tau-\Delta t)\,d\tau \tag{9}$$

with t_s denoting a starting time when the output activity enters another region of the output-input transfer function. However, the solution expression of $\Phi(t)$ in Eq. (9) is dependent on its past history from $t_s - \Delta t$ to $t - \Delta t$. A more practical but narrow form of

the solution can be obtained as follows:

Suppose $\frac{\Phi(t)}{\Phi(t+\Delta t)} = g(\Delta t) = g^{(R)}(\Delta t) + i\, g^{(I)}(\Delta t)$ for $t \geq t_s$. So, the next state $\Phi(t+\Delta t)$ depends only on the present state $\Phi(t)$, and thus the solution of temporal patterns can be expressed either as a discrete time form

$$\Phi(t) = g(\Delta t)^{-\frac{t-t_s}{\Delta t}} \Phi(t_s) \tag{10}$$

when $\frac{t-t_s}{\Delta t}$ is a positive integer, or as a continuous time form

$$\Phi(t) = \exp\left\{\left[K g^{(R)}(\Delta t) + \frac{\rho}{C f'(v_t)} - a + iK g^{(I)}(\Delta t)\right](t-t_s)\right\}\Phi(t_s), \tag{11}$$

from the first of Eq. set (7). Here $g^{(R)}(\Delta t)$ and $g^{(I)}(\Delta t)$ are real and must satisfy:

$$\left.\begin{aligned} g^{(R)}(\Delta t) &= \cos\left[-K\Delta t\, g^{(I)}(\Delta t)\right] \exp\left[\left(a - \tfrac{\rho}{C f'(v_t)}\right)\Delta t\right] \exp\left[-K\Delta t\, g^{(R)}(\Delta t)\right] \\ g^{(I)}(\Delta t) &= \sin\left[-K\Delta t\, g^{(I)}(\Delta t)\right] \exp\left[\left(a - \tfrac{\rho}{C f'(v_t)}\right)\Delta t\right] \exp\left[-K\Delta t\, g^{(R)}(\Delta t)\right] \end{aligned}\right\}, \tag{12}$$

because Eq. (10) and Eq. (11) should be equivalent at $t = t_s + \jmath\Delta t$ for all positive integers $\jmath$. Note that this procedure makes Eq. (10) and Eq. (11) completely equivalent for all time, i.e., Eq. (10) is no longer for discrete time only.

Moreover, from the graph analysis of Eq. set (12), we know the following:

1. The solution for $g^{(I)}(\Delta t)$ in the second of Eq. set (12) always exists, and $g^{(I)}(\Delta t) = 0$ is a trivial solution.

2. Because $g^{(I)} = 0$ is one of the solutions for the second of Eq. set (12), then by the first of Eq. set (12) either $K \geq 0$ or $\frac{1}{\Delta t}, \exp\left[\left(\frac{\rho}{C f'(v_t)} - a\right)\Delta t - 1\right] \geq -K > 0$ guarantees the existence of a solution for $g^{(R)}(\Delta t)$. This gives the existence condition for a real solution $g(\Delta t)$. Neurophysiologically, the time delay Δt is a small interval of the order 1×10^{-3} *sec*, and the above restriction is well satisfied. On the other hand, if there is a solution for $g^{(I)}$ other than $g^{(I)} = 0$, then solutions $g^{(R)}(\Delta t)$ for the first of Eq. set (12) always exist. In this case, $g(\Delta t)$ has complex solutions.

A special case is obtained for the no delay model, i.e., when $\Delta t = 0$. Then from Eq. set (12), $g^{(R)}(\Delta t) = 1$ and $g^{(I)}(\Delta t) = 0$, and the temporal part reduces to

$$\Phi(t) = \exp\left[\left(K + \frac{\rho}{C f'(v_t)} - a\right)(t-t_s)\right]\Phi(t_s)$$

from Eq. (11). Another interesting case is that from Eq. (11) a periodical solution could theoretically exist in one linear region if

$$\left.\begin{array}{r} K g^{(R)}(\Delta t) - a + \frac{\rho}{C f'(v_t)} = 0 \\ g^{(I)}(\Delta t) \neq 0 \end{array}\right\}.$$

Therefore, periodical phenomena are possible for the time evolution of average cell activity. This agrees with the neurobiological oscillations observed for a group of cells [TMW89]. Besides, this result also shows that for the special form of the solution dictated by Eq. (11), a non-zero time-delay factor Δt is requisite if oscillation is to occur just in one linear region.

Because the lateral connection strength is a function of absolute distance between two neurons [OM89], such kernel is not identically zero and is symmetric continuous, i.e., $w(\vec{r},\vec{r}\,') = w(\vec{r}\,',\vec{r})$. Thus, the second of Eq. set (7) becomes an eigenvalue problem and there exists one or more eigenvalues [Smi64]. For convenience, we define a linear functional operator $\hat{W}$ called the conductance operator and rewrite the equation as:

$$\hat{W}\Psi_\ell(\vec{r}) = \frac{C f'(v_t)}{b\varrho} K_\ell \Psi_\ell(\vec{r}) \qquad (\ell = 1, 2, 3, \cdots), \tag{13}$$

where $\Psi_\ell(\vec{r})$ and $\frac{C f'(v_t)}{b\varrho} K_\ell$ are the eigenfunction and eigenvalue respectively associated with the subscript ℓ which denotes the quantum number. Because $\Psi_\ell(\vec{r})$ must be zero on the boundary of $\mathcal{D}$, this Dirichlet boundary condition usually makes K_ℓ discrete. Moreover, $w(\vec{r},\vec{r}\,')$ is real, so $\hat{W}$ is hermitian. Thus, all the corresponding eigenvalues are real and a set of mutually orthonormal eigenfuctions $\Psi_\ell(\vec{r})$'s can always be chosen, such that,

$$\int_{\mathcal{D}} \Psi_\ell(\vec{r}) \Psi^*_{\ell'}(\vec{r})\, d\vec{r} = \delta_{\ell\ell'},$$

where $\delta_{\ell\ell'}$ denotes the Kronecker delta. Here, the superscript "*" means complex conjugate. If $\hat{W}$ is also compact, all of its eigenfunctions form a complete set in a Hilbert space and any spatial pattern belonging to this space can be spanned by these basis functions [PS90].

The quantization of the spatial part forces the corresponding temporal pattern to be also quantized. From Eqs. (9), (10) and (11), we get

$$\left.\begin{array}{rcl} \Phi_\ell(t) & = & \exp\left\{\left[\frac{\rho}{C f'(v_t)} - a\right](t - t_s)\right\}\Phi_\ell(t_s) \\ & + & K_\ell \int_{t_s}^{t} \exp\left\{\left[\frac{\rho}{C f'(v_t)} - a\right](t - \tau)\right\}\Phi_\ell(\tau - \Delta t)d\tau \\ \Phi_\ell(t) & = & g_\ell(\Delta t)^{-\frac{t - t_s}{\Delta t}}\Phi_\ell(t_s) \\ \Phi_\ell(t) & = & \exp\left\{\left[K_\ell\, g_\ell^{(R)}(\Delta t) + \frac{\rho}{C f'(v_t)} - a + i K_\ell\, g_\ell^{(I)}(\Delta t)\right](t - t_s)\right\}\Phi_\ell(t_s) \end{array}\right\}. \tag{14}$$

By the theory of differential equation and Eqs. (6), (8), and (13), the most general solution for the associative spatio-temporal patterns of output activity which are storable in the neural system, is given by:

$$v(\vec{r},t) = \sum_{\ell} p_{\ell}\Phi_{\ell}(t)\Psi_{\ell}(\vec{r}) + \frac{b\varrho R\int_{\mathcal{D}} w(\vec{r},\vec{r}\,')d\vec{r}\,' - \theta + f'(v_t)v_t - f(v_t)}{f'(v_t) - b\varrho R\int_{\mathcal{D}} w(\vec{r},\vec{r}\,')d\vec{r}\,' - R\rho},$$

with $\Phi_{\ell}(t)$ shown in Eq. set (14). Note that all p_{ℓ}s are uniquely determined by the starting pattern $v(\vec{r},t_s)$, because from the property of orthonormality,

$$p_{\ell} = \frac{1}{\Phi_{\ell}(t_s)}\int_{-\infty}^{\infty}\left[v(\vec{r},t_s) - \frac{b\varrho R\int_{\mathcal{D}} w(\vec{r},\vec{r}\,')d\vec{r}\,' - \theta + f'(v_t)v_t - f(v_t)}{f'(v_t) - b\varrho R\int_{\mathcal{D}} w(\vec{r},\vec{r}\,')d\vec{r}\,' - R\rho}\right]\Psi_{\ell}^{*}(\vec{r})\,d\vec{r}.$$

As a result of above analyses, the dynamic behavior of an associative storable pattern is completely predicted after the initial pattern of external input stimulus is given.

3.2 Pattern retrieval process

Retrieval is a process that attempts to recover a degraded pattern by using some *prior knowledge* of the degradation phenomenon. As in many other pattern retrieval schemes, our goal here is to "improve" an initial noisy pattern in some sense. Since biological systems for retrieval use feedback and resonance, we study an iterative rather than a "one-pass" scheme. The mechanism is shown in Fig. (2). Recollect that an external input pattern

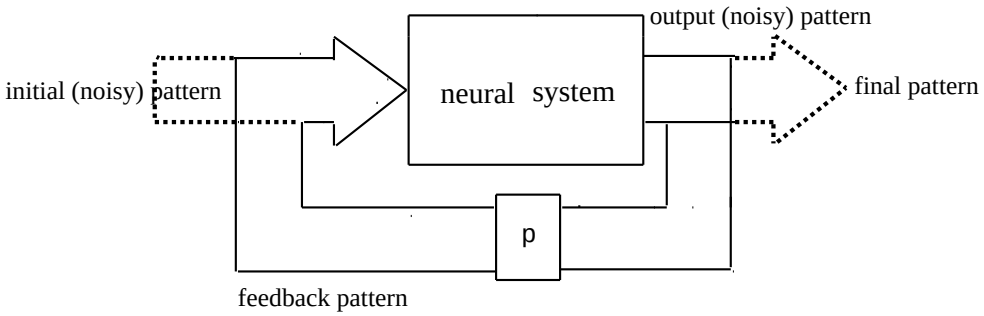

Figure 2: System model of the retrieval process

$I^{(s)}(\vec{r},t)$ is considered associative or storable if $I^{(s)}(\vec{r},t) = \rho v(\vec{r},t+\Delta t)$. If a noisy version of this pattern, expressed as $I^{(s)}(\vec{r},t) + \rho\epsilon^{(0)}(\vec{r},t)$ is presented instead, we expect the output

pattern will be $v(\vec{r}, t+\Delta t)+\epsilon^{(1)}(\vec{r}, t+\Delta t)$. Here, $\epsilon^{(1)}(\vec{r}, t+\Delta t)$ is the corresponding output noise, and the superscripts $^{(0)}$ and $^{(1)}$ denote the zeroth and the first iterations respectively.

In the retrieval process, we receive a noisy pattern (external stimuli) at some particular time t' and make it stay in the dynamical state associated at time t'. If the feedback signal is closer to a stored signal than the original noisy signal, it may be possible to further reduce the noise in subsequent iterations. Step by step, we hope to erase the noise and figure out the original unnoisy signal (one of certain stored patterns). However, it is also obvious from Eq. (5) that whenever the initial noise is expressible as

$$\epsilon^{(0)}(\vec{r}, t) = \sum_{\ell} q_{\ell} \Phi_{\ell}(t+\Delta t) \Psi_{\ell}(\vec{r}), \tag{15}$$

it cannot be reduced further. The reason is that noise of form given in Eq. (15) when added to a storable pattern, translates it to another storable one.

By the notations given above, we rewrite Eq. (5) as a noisy pattern form at that time t':

$$\begin{aligned} & \frac{d\left[v(\vec{r}, t+\Delta t)+\epsilon^{(1)}(\vec{r}, t+\Delta t)\right]}{dt}|_{t=t'} = -a\left[v(\vec{r}, t'+\Delta t)+\epsilon^{(1)}(\vec{r}, t'+\Delta t)\right] \\ + & \frac{b\varrho}{Cf'(v_{t'})}\hat{W}\left[v(\vec{r}, t')+\epsilon^{(1)}(\vec{r}, t')\right]+\frac{\rho}{Cf'(v_t)}\left[v(\vec{r}, t'+\Delta t)+\epsilon^{(0)}(\vec{r}, t')\right] \\ + & \frac{1}{f'(v_t)}\left[\frac{b\varrho}{C}\int_{\mathcal{D}} w(\vec{r}, \vec{r}')d\vec{r}' - a\theta + af'(v_t)\, v_t - af(v_t)\right]. \end{aligned} \tag{16}$$

When there is no noise, Eq. (16) reduces to the pattern-storing equation at time t'. Suppose the noise is not significantly large so that both equations are governed by the same linear section of the neural transfer function. Then the time dependent noise-reducing equation can be derived by taking the difference between Eq. (16) and the pattern-storing equation at time t'. That is

$$\frac{d\left[\epsilon^{(1)}(\vec{r}, t+\Delta t)\right]}{dt}|_{t=t'} = -a\,\epsilon^{(1)}(\vec{r}, t'+\Delta t)+\frac{b\varrho}{Cf'(v_t)}\hat{W}\epsilon^{(1)}(\vec{r}, t')+\frac{\rho}{Cf'(v_t)}\epsilon^{(0)}(\vec{r}, t'). \tag{17}$$

In an iterative manner, $v(\vec{r}, t')+\epsilon^{(1)}(\vec{r}, t')$ becomes a new input feedback pattern and $v(\vec{r}, t')+\epsilon^{(2)}(\vec{r}, t')$ will be the next output pattern. This process is repeated many times and if the noise is reduced to some tolerant value, we say the noisy pattern is retrievable; else it is considered unretrievable.

In experimental observations, usually the input noise is stationary. Thus, we seek a time independent solution to the time independent noise-reducing equation, which is

$$\epsilon^{(1)}(\vec{r}) = \frac{R}{f'(v_t)} \left[\rho \epsilon^{(0)}(\vec{r}) + b\varrho \, \hat{W} \epsilon^{(1)}(\vec{r}) \right], \tag{18}$$

from Eq. (17). The second term in Eq. (18) can be expanded as an absolutely and uniformly convergent series in eigenfunctions of the operator $\hat{W}$ [Smi64]. By finding the expansion coefficients, we can then get the solution of Eq. (18) as follows:

$$\epsilon^{(1)}(\vec{r}) = \frac{R\rho}{f'(v_t)} \sum_{\ell} \left(1 - \frac{K_\ell}{a} \right)^{-1} \varepsilon_\ell^{(0)} \, \Psi_\ell(\vec{r}), \tag{19}$$

with

$$\varepsilon_\ell^{(0)} = \int_{-\infty}^{\infty} \epsilon^{(0)}(\vec{r}) \, \Psi_\ell^*(\vec{r}) \, d\vec{r},$$

i.e., the Fourier coefficients of function $\epsilon^{(0)}(\vec{r})$ with respect to bases $\Psi_\ell(\vec{r})$'s. When some of K_ℓ's are equal to a, these corresponding basis functions must be extracted from the summation and become separate terms. Such a critical case will be omitted from the present consideration.

Recursively inserting $\epsilon^{(\imath-1)}(\vec{r})$ into Eq. (19) to find out $\epsilon^{(\imath)}(\vec{r})$ from $\imath = 1$ to m, the mth iterative noise becomes

$$\begin{aligned} \epsilon^{(m)}(\vec{r}) &= \left[\frac{R\rho}{f'(v_t)} \right]^m \sum_{\ell} \left(1 - \frac{K_\ell}{a} \right)^{-m} \varepsilon_\ell^{(0)} \, \Psi_\ell(\vec{r}) \\ &= \left[\frac{R\rho}{f'(v_t)} \right]^m \sum_{\kappa=0}^{m} C_\kappa^m \left[\sum_{\ell} \left(\frac{K_\ell}{a - K_\ell} \right)^{\kappa} \varepsilon_\ell^{(0)} \, \Psi_\ell(\vec{r}) \right]. \end{aligned} \tag{20}$$

This result implies that

$$\left| \epsilon^{(m)}(\vec{r}) \right| < \left[\frac{2R\rho}{f'(v_t)} \right]^m sup_{(\kappa)} \left| F_{(\kappa)}(\vec{r}) \right|, \tag{21}$$

with

$$F_{(\kappa)}(\vec{r}) = \sum_{\ell} \left(\frac{K_\ell}{a - K_\ell} \right)^{\kappa} \varepsilon_\ell^{(0)} \, \Psi_\ell(\vec{r}) \qquad\qquad 1 \le \kappa \le m. \tag{22}$$

Suppose that $\left[\frac{2R\rho}{f'(v_t)} \right]^m$ converges faster than $\frac{1}{sup_{(\kappa)} \left| F_{(\kappa)}(\vec{r}) \right|}$ for any point $\vec{r} \in \mathcal{R} \subset \mathcal{D}$, then the corresponding pattern $v(\vec{r}, t)$ with $\vec{r} \in \mathcal{R}$ is retrievable and our neural system can be regarded as a pattern retrieval filter.

4 Neural Model of a Visual Cortex

There is both anatomical and physiological evidence from the mammalian brain that the following type of lateral interaction exists between cortical cells [Swi80, Koh82, BG68]:

1. Lateral excitation reaching up to a radius of 50 to 200 μm (in primates).

2. The excitatory area is surrounded by a penumbra of inhibitory action reaching up to a radius of 200 to 600 μm.

3. A weaker excitatory action surrounds the inhibitory penumbra and reaches up to a radius of several centimeters. The degree of lateral interaction is usually described as having the form of a Mexican hat (Fig. 3).

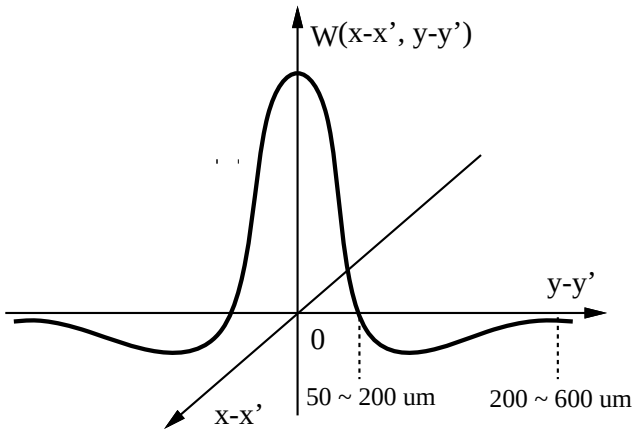

Figure 3: The "Mexican-hat function" of lateral interaction in 2-d

4. In human beings, the total cortical surface area is approximately 1600 cm^2 and more than half of it is related to vision.

These biological evidences suggest that the lateral connections can be regarded as short-range interaction if the longer-range weak excitatory action is neglected. Thus, many researchers have chosen a difference-of-Gaussian function to approximate such interconnections [Ama77, Swi80, HS85]. In this section we choose a similar weight kernel:

$$w(\vec{r},\vec{r}') = A_P \exp\left[-a_P \left|\vec{r}-\vec{r}'\right|^2\right] - A_N \exp\left[-a_N \left|\vec{r}-\vec{r}'\right|^2\right] \tag{23}$$

to perform our analysis. Note that we have given a general $\mathcal{N}$-d expression of the weight kernel. Moreover,

$$\int_{-\infty}^{\infty} w(\vec{r},\vec{r}\,')d\vec{r}\,' = (\sqrt{2}\pi)^{\frac{\mathcal{N}}{2}}\left[A_{_P}a_{_P}^{\frac{-\mathcal{N}}{4}} - A_{_N}a_{_N}^{\frac{-\mathcal{N}}{4}}\right]$$

is a constant, which supports our claim in Section 3.

In Eq. (23), $A_{_P}, a_{_P}A_{_N}$ and $a_{_N}$ can be chosen appropriately to satisfy the required features, i.e. as the distance $|\vec{r}-\vec{r}\,'|$ increases, $w(\vec{r},\vec{r}\,')$ decays to its minimum (a negative value) and then approaches zero monotonically within a short distance. This allows us to replace the integral domain $\mathcal{D}$ with an infinite space for convenience of calculation, when we solve the neural equations later. Because $w(\vec{r},\vec{r}\,')$ approaches to zero in a short distance, the region external to the cortex domain has little contribution to the integration. This approximation releases the Dirichlet boundary condition. As we can see in the following paragraphs, all the formulae of the discrete forms derived in the last section will yield their continuous versions.

As usual, we use a sigmoidal input-output transfer function $v = tanh[\lambda(u-\theta)]$. Then, we obtain its inverse function $u-\theta = f(v) = \frac{1}{2\lambda}ln\left(\frac{1+v}{1-v}\right)$ and derivative function $f'(v) = \frac{1}{\lambda(1-v^2)}$ with respect to v. From Eq.(3), the three linearized sections of the inverse input-output transfer function are given by:

$$f(v(\vec{r},t)) \approx \frac{1}{2\lambda}ln\left(\frac{1+v_t}{1-v_t}\right) + \frac{1}{\lambda(1-v_t^2)}[v(\vec{r},t)-v_t]. \tag{24}$$

Consider a special case of the cortical model of Section 3, which is characterized by Eqs. (23) and (24). Given a quantum number $\mathcal{K}$, the corresponding eigenvalue of the conductance operator $\hat{W}$ defined in Eq. (13) is

$$\frac{CK_{(\mathcal{K})}}{b\varrho\lambda(1-v_t^2)} = \pi^{\frac{\mathcal{N}}{2}}\left[\frac{A_{_P}}{\sqrt{a_{_P}}^{\mathcal{N}}}\exp\left(\frac{-\mathcal{K}}{4a_{_P}}\right) - \frac{A_{_N}}{\sqrt{a_{_N}}^{\mathcal{N}}}\exp\left(\frac{-\mathcal{K}}{4a_{_N}}\right)\right],$$

and its eigenfunction is

$$\Psi_{(\mathcal{K})}(\vec{r}) = (2\pi)^{\frac{-\mathcal{N}}{2}}\sum_{\mu} p_{_\mu}\exp\left(i\,\vec{k_\mu}\cdot\vec{r}\right) \tag{25}$$

with $\vec{k_\mu}\cdot\vec{k_\mu} = \mathcal{K}$. The spatial patterns of Eq. (25) are a generalization of the patterns obtained by Ermentrout and Cowan in 1979 [EC79]. Through an inverse logarithmic mapping, the visual hallucination phenomena mentioned in [EC79, Sch80] can also be explained from this visual cortex model.

There are more than one independent eigenfunctions that corresponds to the same eigenvalue of the conductance operator $\hat{W}$. When this occurs, we have a degeneracy. Mathematically, the orthogonality of eigenfunctions is guaranteed and can be obtained by proper linear combination of these degenerate states. In our visual cortex system, it is easy to see that for given continuous quantum numbers $(k_1, k_2, \cdots, k_{\mathcal{N}}) = \vec{k}$, we have the eigenvalue

$$\frac{C K_{(\vec{k})}}{b\varrho\lambda(1-v_t^2)} = \pi^{\frac{\mathcal{N}}{2}} \left[\frac{A_{_P}}{\sqrt{a_{_P}}\mathcal{N}} \exp\left(\frac{-\vec{k}\cdot\vec{k}}{4a_{_P}}\right) - \frac{A_{_N}}{\sqrt{a_{_N}}\mathcal{N}} \exp\left(\frac{-\vec{k}\cdot\vec{k}}{4a_{_N}}\right) \right]. \tag{26}$$

Its eigenfunction is uniquely determined by

$$\Psi_{(\vec{k})}(\vec{r}) = (2\pi)^{\frac{-N}{2}} \exp\left(i\,\vec{k}\cdot\vec{r}\right), \tag{27}$$

which forms a complete set in an L_2 over $E^{\mathcal{N}}$ space and satisfies the "orthonormal" condition

$$\int_{-\infty}^{\infty} \Psi_{(\vec{k})}(\vec{r})\, \Psi^*_{(\vec{k}\,')}(\vec{r})\, d\vec{r} = \delta\left(\vec{k} - \vec{k}\,'\right), \tag{28}$$

with $\delta\left(\vec{k} - \vec{k}\,'\right)$ the Dirac delta function.

Following the same procedures as in the last section (except that the discrete technique is replaced with the continuous one), we can obtain all the parallel results. According to Eq. sets (12) and (14), three alternative expressions of the temporal pattern $\Phi_{(\vec{k})}(t)$ are:

$$\left.\begin{aligned}
\Phi_{(\vec{k})}(t) &= \exp\left\{\left[\tfrac{\rho\lambda(1-v_t^2)}{C} - a\right](t-t_s)\right\}\Phi_{(\vec{k})}(t_s) \\
&+ K_{(\vec{k})}\int_{t_s}^{t} \exp\left\{\left[\tfrac{\rho\lambda(1-v_t^2)}{C} - a\right](t-\tau)\right\}\Phi_{(\vec{k})}(\tau-\Delta t)d\tau \\
\Phi_{(\vec{k})}(t) &= \left[g^{(R)}_{(\vec{k})}(\Delta t) + i\, g^{(I)}_{(\vec{k})}\right]^{-\frac{t-t_s}{\Delta t}} \Phi_{(\vec{k})}(t_s) \\
\Phi_{(\vec{k})}(t) &= \exp\left\{\left[K_{(\vec{k})} g^{(R)}_{(\vec{k})}(\Delta t) + \tfrac{\rho\lambda(1-v_t^2)}{C} - a + iK_{(\vec{k})} g^{(I)}_{(\vec{k})}(\Delta t)\right](t-t_s)\right\}\Phi_{(\vec{k})}(t_s)
\end{aligned}\right\}, \tag{29}$$

with $g^{(R)}_{(\vec{k})}$ and $g^{(I)}_{(\vec{k})}$ determined by

$$\left.\begin{aligned}
g^{(R)}_{(\vec{k})}(\Delta t) &= \cos\left[-K_{(\vec{k})}\Delta t\, g^{(I)}_{(\vec{k})}(\Delta t)\right] \exp\left[\left(a - \tfrac{\rho\lambda(1-v_t^2)}{C}\right)\Delta t\right] \exp\left[-K_{(\vec{k})}\Delta t\, g^{(R)}_{(\vec{k})}(\Delta t)\right] \\
g^{(I)}_{(\vec{k})}(\Delta t) &= \sin\left[-K_{(\vec{k})}\Delta t\, g^{(I)}_{(\vec{k})}(\Delta t)\right] \exp\left[\left(a - \tfrac{\rho\lambda(1-v_t^2)}{C}\right)\Delta t\right] \exp\left[-K_{(\vec{k})}\Delta t\, g^{(R)}_{(\vec{k})}(\Delta t)\right]
\end{aligned}\right\}. \tag{30}$$

So, a general associative pattern of spatio-temporal output activity for the visual model presented in this section is solved as

$$v(\vec{r}, t) = (2\pi)^{\frac{-\mathcal{N}}{2}} \int_{-\infty}^{\infty} p(\vec{k})\, \Phi_{(\vec{k})}(t) \exp\left(i\,\vec{k}\cdot\vec{r}\right) d\vec{k} + c, \tag{31}$$

with constant

$$c = \frac{\frac{1}{\lambda}\left[\frac{v_t}{1-v_t^2} - \frac{1}{2}ln\left(\frac{1+v_t}{1-v_t}\right)\right] + (\sqrt{2}\pi)^{\frac{\mathcal{N}}{2}} b\varrho R\left(A_P a_P^{\frac{-\mathcal{N}}{4}} - A_N a_N^{\frac{-\mathcal{N}}{4}}\right) - \theta}{\frac{1}{\lambda(1-v_t^2)} - (\sqrt{2}\pi)^{\frac{\mathcal{N}}{2}} b\varrho R\left(A_P a_P^{\frac{-\mathcal{N}}{4}} - A_N a_N^{\frac{-\mathcal{N}}{4}}\right) - R\rho}$$

from Eq. (8). In addition, from Eqs. (28) and (31)

$$p(\vec{k}) = \frac{(2\pi)^{\frac{-\mathcal{N}}{2}}}{\Phi_{(\vec{k})}(t_s)} \int_{-\infty}^{\infty} [v(\vec{r}, t_s) - c] \exp\left(-i\,\vec{k}\cdot\vec{r}\right) d\vec{r},$$

which means that the coefficient $p(\vec{k})$ is uniquely determined by a starting pattern $v(\vec{r}, t_s)$. Thus, the associative spatio-temporal pattern for storing is completely governed by Eq. (31).

An approximation of the sigmoidal function is a linear middle line $v(t) = \lambda\,[u(t) - \theta]$ bounded by two saturation lines $v(t) = -1$ and $v(t) = 1$. The connection points are $\left(\frac{-1}{\lambda}, -1\right)$ and $\left(\frac{1}{\lambda}, 1\right)$. For the middle region, we set $v_t = 0$ in Eqs. (29), (30) and (31) to get the desired solution. For the saturation regions, by setting $v_t \to \pm 1$ we get $\frac{1}{(1-v_t^2)} \to \infty$, $(1 - v_t^2) ln\left(\frac{1+v_t}{1-v_t}\right) \to 0$, $K_{(\vec{k})} \to 0$ and $\Phi_{(\vec{k})}(t_s) \to 0$. Thus, the spatio-temporal output activity in these extreme regions is nothing but

$$v(\vec{r}, t) = v_t \to \pm 1,$$

which is as expected. Therefore, all the information is embodied in the middle region. For instance, if $-1 < v(\vec{r_0}, t_s) < 1$ for a particular position $\vec{r_0}$, it is in the middle region at the starting time, t_s. $\Phi_{(\vec{k})}(t)$ evolves with time, and finally $v(\vec{r_0}, t)$ may reach a critical value (-1 or 1). Although the system variable $v(\vec{r_0}, t)$ then enters another region of the output-input transfer function, the same equation as has been used in the middle section can equivalently determine the behavior of $v(\vec{r_0}, t)$ in the extending region which includes the boundary point -1 or 1.

Next, let us consider the pattern retrieval capability of the visual system. The time independent noise-reducing equation is

$$\epsilon^{(\iota)}(\vec{r}) = \lambda R\left(1 - v_t^2\right)\left[\rho\epsilon^{(\iota-1)}(\vec{r}) + b\varrho\hat{W}\epsilon^{(m)}(\vec{r})\right]$$

and its corresponding mth iterative result, which is similar to Eq. (20), is

$$\epsilon^{(m)}(\vec{r}) = (2\pi)^{\frac{-\mathcal{N}}{2}}\left[R\rho\lambda(1 - v_t^2)\right]^m \int_{-\infty}^{\infty}\left(1 - \frac{K_{(\vec{k})}}{a}\right)^{-m} \varepsilon^{(0)}(\vec{k}) \exp\left(i\,\vec{k}\cdot\vec{r}\right) d\vec{k} \qquad (32)$$

$$= \quad (2\pi)^{\frac{-\mathcal{N}}{2}} \left[R\rho\lambda(1-v_t^2)\right]^m \sum_{\kappa=0}^{m} C_{\kappa}^{m} \left[\int_{-\infty}^{\infty} \left(\frac{K_{(\vec{k})}}{a - K_{(\vec{k})}}\right)^{\kappa} \varepsilon^{(0)}(\vec{k}) \exp\left(i\,\vec{k}\cdot\vec{r}\right) d\vec{k}\right],$$

with

$$\varepsilon^{(0)}(\vec{k}) = (2\pi)^{\frac{-\mathcal{N}}{2}} \int_{-\infty}^{\infty} \epsilon^{(0)}(\vec{r}) \exp\left(-i\,\vec{k}\cdot\vec{r}\right) d\vec{r}$$

i.e., the Fourier transformation of $\epsilon^{(0)}(\vec{r})$. Similar to Eqs. (21) and (22), we get

$$\left|\epsilon^{(m)}(\vec{r})\right| < \left[2R\rho\lambda(1-v_t^2)\right]^m sup_{(\kappa)} \left|F_{(\kappa)}(\vec{r})\right|,$$

with

$$F_{(\kappa)}(\vec{r}) = (2\pi)^{\frac{-\mathcal{N}}{2}} \int_{\infty}^{\infty} \int_{\infty}^{\infty} \left(\frac{K_{(\vec{k})}}{a - K_{(\vec{k})}}\right)^{\kappa} \varepsilon^{(0)}(\vec{k}) \exp\left(i\,\vec{k}\cdot\vec{r}\right) d\vec{k}.$$

5 Application to invariant pattern recognition

Invariance with respect to certain transformations is one of the main characteristics of a versatile pattern-recognition system [BC91]. One considers a neural visual model to be incomplete if it is not able to identify an object from its certain transformed versions. For example, an object should be identifiable regardless of its location, and the aspect and size of the image projected from it through the retina to the visual cortex. In the following paragraphs, we will apply the noise-reducing equation introduced in the previous section to identify some invariant features of the visual cortex model developed in this paper.

Most neural networks in the brain, especially those in the cerebral neocortex, are essentially 2-d layers of processing units (cell or cellular modules) [Koh82]. Let us pay attention to a 2-d pattern in this section. The approach can be easily generalized to an $\mathcal{N}$-d case in a similar way.

A general 2-d affine transformation formulation

$$\left.\begin{aligned} x^{(old)} &= c_{11}\,x^{(new)} + c_{12}\,y^{(new)} + c_{13} \\ y^{(old)} &= c_{21}\,x^{(new)} + c_{22}\,y^{(new)} + c_{23} \end{aligned}\right\}$$

is able to perform transforms of scaling, translation, rotation and mirror-reflection by choosing the coefficients c_{mn} properly [Bar88].

If we pick up a pattern at some time t', then from Eq. (31) the dynamical pattern associated at that time can be expressed as

$$v(x,y,t') = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(k_1,k_2)\Phi_{(k_1,k_2)}(t')\exp\left[i\left(k_1x+k_2y\right)\right]\,dk_1\,dk_2 + c, \qquad (33)$$

Then, a stationary noise

$$\begin{aligned}\epsilon^{(0)}(x,y) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s(k_1,k_2)\exp\left\{i\left[k_1\left(c_{11}x+c_{12}y+c_{13}\right)+k_2\left(c_{21}x+c_{22}y+c_{23}\right)\right]\right\}\,dk_1\,dk_2 \\ &- \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} s(k_1,k_2)\exp\left[i\left(k_1x+k_2y\right)\right]\,dk_1\,dk_2\end{aligned}$$

is the noise which induces the above mentioned transformations of pattern $v(x,y,t)$ at time t'. Here, $s(k_1,k_2) = p(k_1,k_2)\Phi_{(k_1,k_2)}(t')$ for the fixed value t' and theoretically, we take $s(k_1,k_2)$ to be of class L_1 over the space defined by $-\infty < k_1 < \infty$ and $-\infty < k_2 < \infty$ [BC49]. By change of variables in the integrals and some algebraic manipulation, we get the corresponding Fourier transformation of $\epsilon^{(0)}(x,y)$ as

$$\varepsilon^{(0)}(k_1,k_2) = |\Lambda|\exp\left(i\varsigma_3\right)s\left(\varsigma_1,\varsigma_2\right) - s\left(k_1,k_2\right),$$

with $\Lambda = \frac{1}{c_{11}c_{22}-c_{12}c_{21}}$, $\varsigma_1 = \Lambda(k_1c_{22}-k_2c_{21})$, $\varsigma_2 = \Lambda(k_2c_{11}-k_1c_{12})$ and $\varsigma_3 = \varsigma_1c_{13}+\varsigma_2c_{23}$. So,

$$\left|\varepsilon^{(0)}(k_1,k_2)\right| < \left|\Lambda s\left(\varsigma_1,\varsigma_2\right)\right| + \left|s\left(k_1,k_2\right)\right|. \qquad (34)$$

By Eq. (34) and taking absolute values on both sides of Eq. (32), an inequality formula is obtained

$$\left|\epsilon^{(m)}(x,y)\right| < \frac{\left|\Lambda^2+1\right|}{2\pi}\,Max_{(k_1,k_2)}\left|\frac{\rho\lambda(1-v_t^2)}{\frac{1}{R}-CK_{(k_1,k_2)}}\right|^m\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|s(k_1,k_2)\right|\,dk_1\,dk_2. \qquad (35)$$

Obviously, the behavior of inequality (35) depends critically on the electrical characterization of its biological elements. Let us first identify the magnitude of $K_{(k_1,k_2)}$. From Eq. (26), we get the extremum values of $K_{(k_1,k_2)}$,

$$K_{(0,0)} = \frac{\pi b\varrho\lambda(1-v_t^2)}{C}\left(\frac{A_P}{a_P}-\frac{A_N}{a_N}\right) \qquad (36)$$

and

$$K_{(k_x,k_y)} = \frac{\pi b\varrho\lambda(1-v_t^2)}{C}\left[\frac{A_P}{a_P}\exp\left(-a_N\mathcal{K}^{(ext)}\right)-\frac{A_N}{a_N}\exp\left(-a_P\mathcal{K}^{(ext)}\right)\right] \qquad (37)$$

when $k_x^2 + k_y^2 = 4a_{_P}a_{_N}\mathcal{K}^{(ext)}$, with $\mathcal{K}^{(ext)} = \frac{ln\left(\frac{A_P}{A_N}\right)+2\,ln\left(\frac{a_P}{a_N}\right)}{a_{_P}-a_{_N}}$. We know that b is around 10^2 and the surface density of the neuron number ϱ is around 10^5 per mm^2 in area 17 of a cat [BC83, TPW91]. Moreover, the orders of $A_{_P}$ and $A_{_N}$ are 0.1 to 0.01 μS [WDKE89], and the short range of excitation connection is from the cell center to a distance of 200 μm, and inhibition connection is from 200 to 600 μm. With these in mind, we may take $A_{_P} = 0.080\mu S, A_{_N} = 0.036\,\mu S, a_{_P} = 0.8 \times 10^{-4}$ per μm^2 and $a_{_N} = 0.6 \times 10^{-4}$ per μm^2 as values that typically satisfy the above biological data. In the saturation regions, $v_t^2 \to 1$ and retrieval is trivial. In the middle region, λ is the slope of input-output transfer function at $(u-\theta, v) = (0,0)$, which cannot be too large and v_t is in domain $(-1,1)$. We may choose $\lambda = 1$ and $v_t = 0$. Because $\frac{A_P}{a_{_P}} > \frac{A_N}{a_{_N}} > 0$ and $a_{_P} > a_{_N} > 0$, from Eqs. (36) and (37) we know $K_{(k_1,k_2)}$ is always positive for any k_1 and k_2, and the minimum value of $CK_{(k_1,k_2)}$ is 431.91 μS. The leaking conductance $\frac{1}{R}$ is around 0.1 μS [WDKE89] and ρ is an adjustable parameter. Biologically, the dimensionless constant $0 < \frac{\rho\lambda(1-v_t^2)}{CK_{(k_1,k_2)}-\frac{1}{R}} < 1$ is obvious, and thus the *Max* term in Eq. (35) must be smaller when m becomes larger. Because $s(k_1,k_2)$ belongs to class L_1, the integral term is bounded. Therefore, the noise will be reduced to a tolerant value and the pattern can be identified. This theory explains the invariances of scaling, translation, rotation and mirror-reflection transformations in the visual system.

6 Concluding Remarks

This paper provides, for the first time, a rigorous application of a neural field model for pattern association and retrieval. We are able to obtain the exact solution of the storable spatio-temporal patterns for the given visual cortex model. Moreover, the conditions that make the noisy versions of these associative patterns retrievable are determined through a simple pattern-iterative scheme. Interestingly, the proposed model also demonstrates the invariance of certain patterns to affine transformations.

The biological interpretation and mathematical technique presented in this paper represent initial steps in analyzing high-density neural nets. Possible extensions of this work include studying the convergent rate of the retrieval mechanism, and the effects of weight learning on the storable patterns and on the self-organization of the entire neural sys-

tem. Also, the dynamics of more complex systems that contain control sub-networks need to be investigated and compared with extensive simulation and biological experiments [WML90, TPW91, TPW92]. The theory proposed in this paper can potentially be used to construct a robust, biologically motivated system for perception.

Acknowledgements

The authors wish to thank Dr. K.-S. Huang, Mr. K. Liano, Mr. D.-J. Huang, Dr. B. Willner and Dr. W.L. Miranker for their most valuable discussions. We owe special thanks to Dr. R.E. Wyatt, Dr. B.F. Womack and Mr. K.-Y. Chao for their constant encouragement and useful comments.

References

[AA77] S.I. Amari and M. A. Arbib. Competition and cooperation in neural nets. In J. Metzler, editor, *Systems Neuroscience*, pages 119–165. New York, 1977.

[Ama72a] S.I. Amari. Characteristics of random nets of analog neuron-like elements. *IEEE Trans. on System, Man, and Cybernetics*, 2:1443–1463, 1972.

[Ama72b] S.I. Amari. Learning patterns and pattern sequences by self-organizing nets of threshold elements. *IEEE Trans. Comput.*, 21:1197, 1972.

[Ama75] S.I. Amari. Homogeneous nets of neuron-like elements. *Biol. Cybernetics*, 17:211–220, 1975.

[Ama77] S.I. Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybernetics*, 27:77–87, 1977.

[And68] J.A. Anderson. A memory storage model utilizing spatial correlation function. *Kybernetik*, 5:113–119, 1968.

[And70] J. A. Anderson. Two models for memory organization using interacting traces. *Mathematical Biosciences*, 14:137–160, 1970.

[And72] J.A. Anderson. A simple neural network generating an interactive memory. *Math. Bioscience*, 14:197–200, 1972.

[Bar88] M. Barnsley. *Fractals Everywhere*. Academic, 1988.

[BC49] S. Bochner and K. Chandrasekharan. *Fourier Transforms*. Princeton, 1949.

[BC83] C. Beaulieu and M. Colonnier. The number of neurons in the different laminae of the binocular and monocular regions of area 17 in the cat. *J. Comp. Neurol.*, 217:337–344, 1983.

[BC85] C. Beaulieu and M. Colonnier. A laminar analysis of round-asymmetrical and flat-asymmetrical synapses on spins, dendritic trunks and cell bodies in area 17 of the cat. *J. Comp. Neurol.*, 231:180–189, 1985.

[BC91] E. Barnard and D. Casasent. Invariance and neural nets. *IEEE Transactions on Neural Networks*, 2:498–508, 1991.

[Beu56] R.L. Beurle. Properties of a mass of cells capable of regenerating pulses. *Trans. Roy. Soc. London*, B 240:55–94, 1956.

[BG68] S.M. Blinkov and I.I. Glezer. *The human brain in figures and tables*. New York, 1968.

[Boy75] C.C. Boylls. A theory of cerebellar function with applications to locomotion, 1. the physiological role of climbing fiber inputs in anterior lobe operation. Technical Report COINS TR 75C-6, Univ. Mass. at Amherst, 1975.

[Cai61] E.R. Caianiello. Outline of a theory of thought-processes and thinking machines. *J. Theoret. Biol.*, 2:204–235, 1961.

[CGL91] H.-J. Chang, J. Ghosh, and K. Liano. A macroscopic model of neural ensembles 1: Learning-induced oscillations in a cell assembly. Technical Report TR-92-03-78, U.T. Austin, October, 1991.

[Col71] B.P. Coleman. Mathematical theory of lateral sensory inhibition. *Arch. Rat. Mach. Analys.*, 43:79–100, 1971.

[EC79] G.B. Ermentrout and J.D. Cowan. A mathematical theory of visual hallucination patterns. *Biol. Cybernetics*, 34:137–150, 1979.

[EC80] G.B. Ermentrout and J.D. Cowan. Large scale spatially organized activity in neural nets. *SIAM J. Appl. Math.*, 38:1–21, 1980.

[Ede87] G.M. Edelman. *Neural Darwinism.* New York: Basic Books, 1987.

[EG75] S.A. Ellias and S. Grossberg. Pattern formation, contrast control and oscillations in the short-term memory of shunting on-center off-surround networks. *Biol. Cybernetics*, 20:69–98, 1975.

[Erm86] B. Ermentrout. A model for patterns based on neural activity. *The Veliger*, 28:369–388, 1986.

[FC61] B.G. Farley and W.A. Clark. Activity in networks of neural-like elements. In C. Cherry, editor, *Proc. 4th London Symp. on Inf. Theory.* London: Butterworths, 1961.

[Fuk80] K. Fukushima. Neocognitron: A self-organization neural network model for a mechanism of pattern recognition unaffected by a shift in position. *Biol. Cybernetics*, 36:192–202, 1980.

[Gri63] J.S. Griffith. A field theory of neural nets 1. *Bull. Math. Biophys.*, 25:111–120, 1963.

[Gri65] J.S. Griffith. A field theory of neural nets 2. *Bull. Math. Biophys.*, 27:187–195, 1965.

[HA81] G.E. Hinton and J.A. Anderson. *Parallel Models of Associative Memory.* Erlbaum, Hillsdale, NJ, 1981.

[Hop82] J.J. Hopfield. Neural networks and physical systems with emergent selective computational abilities. *Proc. Natl. Acad. Sci. USA*, 79:2554–2558, 1982.

[Hop84] J.J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences, USA*, 81:3058–3092, 1984.

[HS85] S. Hochstein and H. Spitzer. One, few, infinity: linear and nonlinear processing in the visual cortex. In D. Rose and V. G. Dobson, editors, *Models of the Visual Cortex*, pages 341–350. John Wiley and Sons Ltd, 1985.

[Hua90] K.-S. Huang. *Binary Image Algebra: Development (chapter 4 in A digital optical cellular image processor)*. World Scientific Pulbishing Co., 1990.

[Kan88] K.-I. Kanatani. Transformation of optical flow by camera rotation. *IEEE Trans. Pattern Anal. Machine Intell.*, 10:131–143, 1988.

[Koh72] T. Kohonen. Correlation matrix memories. *IEEE Trans. on Comput.*, C-21:353–359, 1972.

[Koh82] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biol. Cybernetics*, 43:59–69, 1982.

[Koh89] T. Kohonen. *Self-Organization and Associative Memory*. Springer-Verlag, 1989.

[KS88] D. Kleinfeld and H. Sompolinsky. Associative neural network model for the generation of temporal patterns: theory and application to central pattern generators. *Biophys. J.*, 54:1039–1051, 1988.

[OM89] G.F. Oster and J.D. Murray. Pattern formation models and developmental constraints. In *J. P. Trinkaus Anniversary*, 1989.

[PS90] D. Porter and D.S.G. Stirling. *Integral Equations*. Cambridge, 1990.

[RHW86] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning Internal Representations by Error Propagation. In J.A. Feldman, P.J. Hayes, and D.E. Rumelhart, editors, *Parallel Distributed Processing*, pages 318–362. MIT Press, 1986.

[Sch80] E.L. Schwartz. Computational anatomy and functional architecture of striate cortex: A spatial mapping aproach to perceptual coding. *Vision Research*, 20:645–669, 1980.

[Smi64] V.I. Smirnov. *Integral Equations and Partial Differential Equations*. Addison-Wesley, 1964.

[Sta76] J.C. Stanley. Simulation studies of a temporal sequence memory model. *Biol. Cybernetics*, 24:121–137, 1976.

[Swi80] N.V. Swindale. A model for the formation of ocular domiance stripes. *Proc. Roy. Soc. Lond.*, B208:243–264, 1980.

[TMW89] R.D. Traub, R. Miles, and R.K.S. Wong. Model of the origin of rhythmic population oscillations in the hippocampal slice. *Science*, 243:1319–1325, 1989.

[TPW91] E. Thomas, P. Patton, and R.E. Wyatt. A computational model of the vertical anatomical organization of primary visual cortex. *Biol. Cybernetics*, 65:189–202, 1991.

[TPW92] E. Thomas, P. Patton, and R.E. Wyatt. A computational model of vertical signal propagation in the primary visual cortex. *to appear in Biol. Cybernetics*, 1992.

[WC73] H.R. Wilson and J.D. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13:55–80, 1973.

[WCF92] B. Willner, H.J. Chang, and J. Foy. Willard graphs. *to appear in Computing*, 47, 1992.

[WDKE89] U. Wehmeier, D. Dong, C. Koch, and D. Van Essen. Modeling the mmmalian visual system. In C. Koch and I. Segev, editors, *Methods in Neuronal Modeling, From Synapses to Networks*, pages 335–359. MIT Press, 1989.

[WHHL89] A. Waibel, T. Hanazawa, G. Hinton, and K.J. Lang. Phoneme recognition using time-delay neural networks. *IEEE Trans. Acoust., Speech, Signal Process.*, 37:328–339, 1989.

[WML90] B.E. Willner, W.L. Miranker, and C.P. Lu. Neural organization of the locomotive oscillator. Technical Report RC 16390 (No.72765), IBM Research Report, T.J. Watson Research Center, 1990.

[WR46] N. Wiener and A. Rosenblueth. The mathematical formation of the problem of conduction of impulses in a network of connected excitable element, specifically in cardiac muscle. *Arch. Inst. Cardiol. (Mexico)*, 16:204–265, 1946.