

Invariant Pattern Recognition by Means of Fast Synaptic Plasticity

JOACHIM BUHMANN AND KLAUS SCHULTEN

*Physik-Department, Technische Universität München,
James-Franck-Str., 8046 Garching, Fed. Rep. Germany ‡*

A two layer neural system for shift-invariant pattern recognition is proposed. Model neurons are endowed with "physiological dynamics" involving membrane potentials and axonic spikes. Synapses between the two layers are plastic and change according to spike coincidences (Hebbian rules). The first neural network (encoder network) extracts features from a presented pattern and codes the neighborhood relationship of features by coincident activity of neurons. The second network (memory network) has stored several patterns. During recognition of a presented pattern the neural system establishes a strong projection between the first and the second layer, enhances activity in the set of those neurons, which represent the presented patterns, and suppresses activity of other neurons. Synaptic plasticity according to Hebbian rules allows to generate a projection which preserves feature neighborhood relationships. The recognition system is designed according to suggestions of v.d. Malsburg (1981).

1. Introduction

Neural modeling has attracted much attention in the scientific community. Cognitive science and artificial intelligence are beginning to embrace dynamical models like neural networks for information processing. These models share some properties of biological networks, such as fault tolerance, massive parallelism without a system-wide clock cycle and high connectivity and can contribute to our understanding of collective computation in systems with a large number of simple elements. The Hopfield model [1] and related networks, in particular, have provided insight in dynamics of associative memories. In this paper we address the problem of shift-invariant recognition of patterns where neural modeling has shown promising preliminary results [2]. The invariance problem is one of the central problems of pattern recognition and still awaits a satisfactory solution.

Pattern recognition can be divided into two different problems – the problem of encoding a presented pattern invariant of its size, position and slight distortions of its form, and the problem to compare an encoded pattern with several stored pattern. Associative memories can only tackle the second task. They have the ability to search in the space of Hamming distances, i.e. to look for that pattern stored which is most similar to the presented one. Therefore, it is necessary to normalize a presented pattern, i.e. to shift, to rotate and to change the size of a pattern, before the associative abilities of distributed memories can be evoked.

Invariant encoding of a presented pattern can be achieved by separation of absolute information on pattern position, size or form from internal relations between different

‡ Electronic mail addresses: buhmann@dgatum5p.bitnet, kschulte@dgatum5p.bitnet

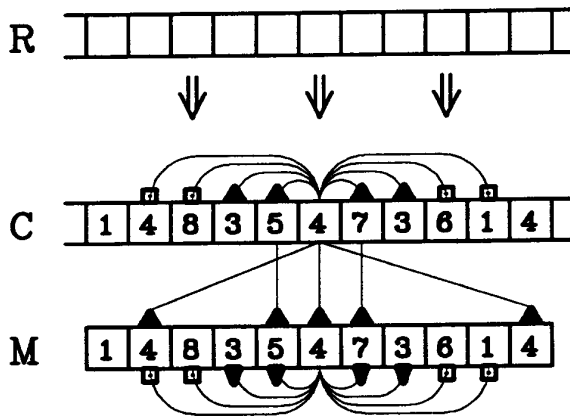


Fig. 1. Schematical description of the neural system designed for invariant pattern recognition. Information flows from model retina R to encoding layer C which extracts features from the presented pattern. Layer M plays the role of an associative memory which has stored several patterns. Black triangles indicate excitatory synapses, small white squares show inhibitory synapses. Large squares indicate neurons sensitive for different features denoted by numbers (labels).

parts of the pattern. A concept for invariant pattern recognition has been proposed by v.d. Malsburg [3]. He suggests that neighborhood relationships are encoded by correlated neural spikes. Neurons signalling spatially distant parts of a pattern do not show coincidences in their firing behaviour. Coincident firing of neurons indicates that these neurons extract neighboring features from a pattern.

2. Layered neural system for invariant pattern recognition

The neural system we are studying consists of two neural layers and a retina (R). Schematically, the system is shown in Fig. 1. The model retina serves as an input channel of the neural system. Patterns appear at an arbitrary position on the retina and will be transferred to the first neural network, the encoding layer (C). The retina is much larger than the presented patterns. Therefore, it is necessary to recognize the patterns independent of their actual position. In computer simulations of the neural system we restrict ourselves to the case of one-dimensional patterns. The model discussed can be generalized to two-dimensional patterns.

A first neural network – the encoding network (C) – extracts various features from the presented pattern. The C-layer has the same size as the retina. Reduction of resolution which normally occur during feature extraction is not considered. We separate the problem of feature extraction from the problem of invariant recognition. A further simplification is introduced by the assumption that exactly one feature is extracted at each site of the pattern. Neurons which represent these features are shown in Fig. 1 by small squares containing a label for the feature for which the cell is sensitive. Activity of a labeled neuron at position x indicates that the corresponding feature is present at x .

A second network (M) which serves as associative memory has stored several patterns. These patterns are represented by sets of neurons also called “neural assemblies” [4]. One of the patterns is presented to the system and should be recognized. In Figure 1 layer M contains a pattern which is activated in layer C. By correlated firing of small clusters of neighboring neurons in both layers a presented pattern can be recognized and distinguished from other stored patterns.

3. Synaptic connectivity

The synaptic connectivity from retina **R** to layer **C** is a one-to-one projection. The synapse connecting receptor j with neuron i of the **C**-layer has the strength $R_{ij} = R \delta_{ij}$. This simple connectivity neglects the question how certain features are extracted from a presented pattern. In our model the projection from retina to **C**-layer has only the function to transmit receptor activity and to activate corresponding neurons in layer **C**.

Synaptic interactions between neurons of layer **C** are organized in form of “*near excitation, far inhibition*”. Neighboring cells are excited by a firing neuron, more distant neurons are inhibited. The synaptic strengths have values

$$W_{ik}^{CC} = \begin{cases} W_e & \text{if } |i - k| \leq L_e \\ -W_i & \text{if } L_e < |i - k| \leq L_i \\ -\zeta W_i & \text{if } |i - k| > L_i \end{cases} \quad (1)$$

with range of excitation $L_e = 2$ and range of inhibition $L_i = 4$. Synaptic strengths are chosen as $W_e = 1$, $W_i = -1$; the ratio between near and far inhibition has the value $\zeta = 0.1$. The center-surround-connectivity brings about small clusters of neurons of size $2L_e$ which fire nearly simultaneously. Thereby, neighborhood relations $|i - k| \leq 2L_e$ between two neurons i , k are encoded in simultaneous spikes. Layer **C** with its center-surround-organization encodes a pattern specific feature neighborhood relationships (feature topology). This encoding allows to identify a presented pattern with a pattern stored in a shift-invariant format, i.e. in a format which solely represents the feature topology of a pattern.

Synapses between neurons of layer **M** which belong to one pattern have also a center-surround organization, i.e. the synaptic strengths assume values also given by (1). As in layer **C**, firing neurons of layer **M** encode their neighborhood relations in simultaneous spikes. Synapses between neurons belonging to different patterns are inhibitory, i.e. $W_{ik} = -\lambda W_i$ if $i \in \{\text{pattern } \nu\} \wedge k \in \{\text{pattern } \mu\} \wedge \nu \neq \mu$ with $\lambda = 0.1$. Inhibition between different patterns yields competition between different neural sets. The assembly with the feature topology which matches the input in **C** has the largest probability to fire with high spike rate; the activity of alternative representing different feature topologies is suppressed. assemblies. Synapses inside layer **C** and layer **M** possess no plasticity.

With fast plastic synapses between layer **C** and layer **M** the correct classification of a presented pattern as one of the stored patterns is achieved. These synapses are excitatory and evolve in time according to Hebbian rules [4]. At the beginning of a recognition task inter-layer synapses are weak and preserve feature labels, i.e. there exist only synapses between neurons sensitive for the same feature types F_i . The initial synaptic strengths are

$$W_{ij}^{CM}(0) = W^{CM} \delta_{F_i, F_j}. \quad (2)$$

Inter-layer connection between an activated set of neurons in layer **C** and two sets of neurons in layer **M** are shown in Fig. 2. In the first case (left picture), neurons in **C**

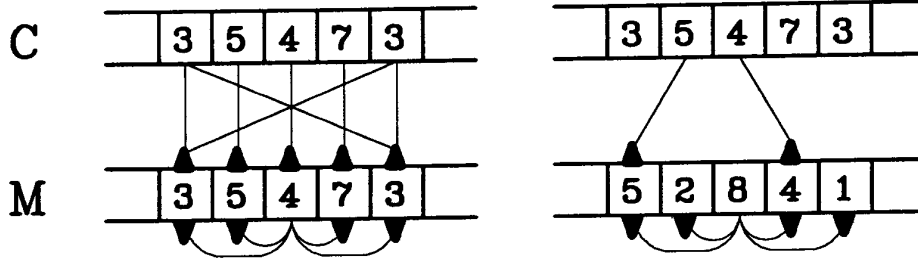


Fig. 2. Comparison of two connectivity structures between layer C and layer M. Excitatory synapses exist between all neurons representing the same feature type.

and in M have identical feature topologies. Therefore, there exists initially an ordered, topology conserving projection between both layers. Activity of a cluster of neurons in layer C yields strong excitation of those neurons in layer M encoding the same part of the pattern. These neurons will fire with high probability. In the second case (right picture) the inter-layer synapses connect two sets of neurons in C and in M, respectively, with different feature topology. There are initially only isolated synapses, but no ordered projection.

4. Dynamics of Membran Potentials

The dynamics of neurons are described in our model by two variables, membran potentials and axonic spikes. The dynamics of the membrane potentials involves two processes, relaxation of the membrane potential and neural interaction as determined by the somatic integration rule. Axonal spikes are generated whenever the membrane potential reaches a threshold value. A detailed description of the neural dynamics is given in [5]. Postsynaptic excitation by presynaptic spikes is described by an exponential activity function with decay time $T_U = 1ms$

$$G_k(\Delta t_k/\tau) = \exp\left(-\frac{\Delta t_k}{\tau}\right), \quad \text{with} \quad \Delta t_k = t - t_k^0. \quad (3)$$

$\Delta t_k = t - t_{0k}$ measures the time elapsed since the last spike of neuron k at t_k^0 . The kinetic equations of the membrane potentials $U_i(t)$ which also include the stochastic fluctuations are given by a system of non-linear coupled Langevin equations

$$\frac{dU_i^L}{dt} = -\frac{U_i^L}{T_R} + \rho[\Delta t_i] \left(\omega^L \sigma[A_i^L(t)] + \frac{\eta^L}{\sqrt{T_R/2}} \xi(t) \right) \quad L \in \{C, M\}. \quad (4)$$

Upper index L denotes the layer of neuron i . The first term in (4) approximates the relaxation of the membrane potential $U_i(t)$ to its resting value $U_0 = 0mV$ within a time interval $T_R = 2.5ms$. The second term in (2) describes the communication of the postsynaptic cell i with the connected neurons and receptors and adds a Gaussian white noise $\xi(t)$ with strength $\eta/\sqrt{T_R/2}$. Noise in an isolated neuron produces a Gaussian

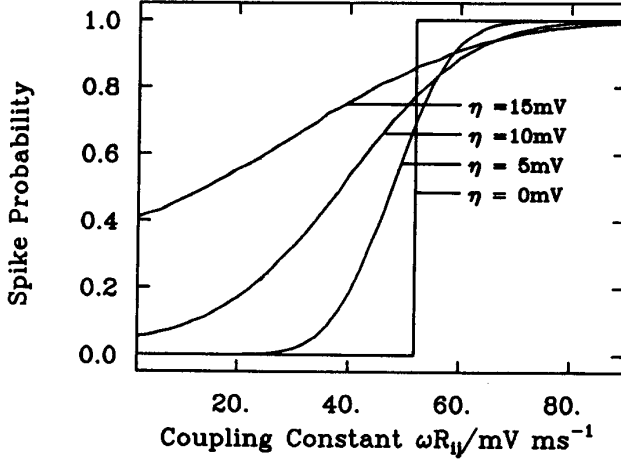


Fig. 3.

Dependence of the probability, that neuron i responds with a spike if receptor j fires, on the coupling constant ωR_{ij} for four different strengths η of noise.

distribution of the membrane potential $U_i(t)$ with mean value $U_0 = 0$ mV and variance $\eta = 10$ mV. Afferent activity $A_i^C(t)$ of a neuron i in layer C consists of receptor activity and of interaction between neurons of layer C, i.e.

$$A_i^C(t) = \sum_j R_{ij} G_j^R \left(\frac{\Delta t_j^R}{T_U} \right) + \sum_k W_{ik}^{CC} G_k^C \left(\frac{\Delta t_k}{T_U} \right) \quad (5)$$

with $i, k \in C$ and $j \in R$. Contributions to the postsynaptic potential of neurons in layer M are due to the interactions between the two layers and due to the interactions within layer M, i.e.

$$A_i^M(t) = \sum_j W_{ij}^{CM} G_j^C \left(\frac{\Delta t_j}{T_U} \right) + \sum_k W_{ik}^{MM} G_k^M \left(\frac{\Delta t_k}{T_U} \right) \quad (6)$$

with $i, k \in M$ and $j \in C$. The coupling constant ω^L and the strength of noise η^L depend on the layer index $L \in \{C, M\}$. Typical values are $\omega^C = 50$ mV/ms, $\omega^M = 100$ mV/ms and $\eta^C = \eta^M = 10$ mV. The sigmoidal function $\sigma[A_i(t)]$ with a linear behaviour for small $A_i(t)$ and a saturation value for strong activity prevents potential changes which are unphysiologically large. Total and relative refractory periods are taken into account by the function $\rho[\Delta t_i]$ which suppresses the sensitivity of neuron i to afferent excitation during a total refractory period $T_F = 5$ ms. The function also lets the neuron gradually regain its sensitivity to incoming excitation or inhibition during a relative refractory period of 5 ms.

Continuous time evolution of potentials in our model is interrupted when the neuron reaches the threshold $U_T = 30$ mV and fires a spike. Instantaneously, the membrane potential is set to a value normally distributed around the refractory potential $U_F = -15$ mV. In this event the time of the last spike t_i^0 is updated and the memory function $G_i(\Delta t_i/T_U)$ is set to the value 1. This behaviour is represented as follows:

$$\text{if } U_i(t) \geq U_T, \text{ then } \begin{cases} t_i^0 = t, \\ U_i(t) = U_F + \xi(t). \end{cases} \quad (7)$$

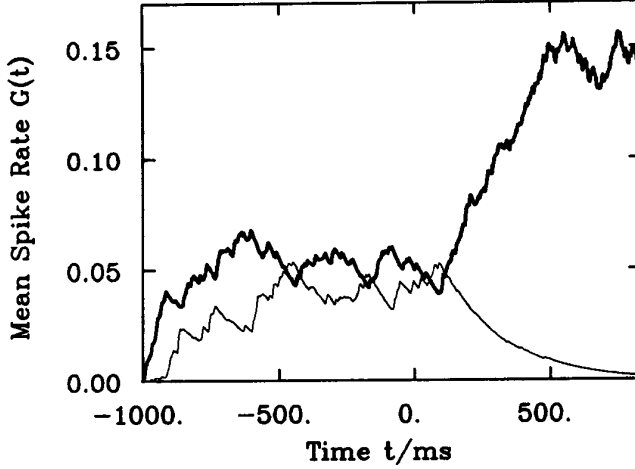


Fig. 4. Mean activity of the two neural assemblies in layer M representing the presented pattern (bold curve) and the second pattern (thin curve). Synapses can change at $t \geq 0$.

The reaction of a neuron to a receptor input depends on the coupling constant ω^L (we assume $R_{ij} = 1$). In case of strong coupling the excited neuron will always reach the threshold whereas weak coupling causes only small postsynaptic potentials which never reach the threshold. Figure 3 shows the probability that a neuron which received a receptor spike at $t = 0$ ms will fire within 5 ms. This probability is presented as a function of the coupling strength ωR_{ik} for four different noise levels ($\eta = 0, 5, 10, 15$ mV).

5. Plasticity of interlayer synapses

In our neural system for invariant pattern recognition we introduce synaptic plasticity on a time scale of 0.2–0.5 s [5]. According to the Hebbian rules the synaptic dynamics is assumed to depend on synchronicity or asynchronicity of the pre- and postsynaptic spikes. In addition to the Hebbian rules we require for synaptic modifications that the mean spike rate of both neurons considerably exceed the spontaneous spike rate $\nu_s \approx 5s^{-1}$. In the case of synchronous firing both neurons satisfy this condition and the synapse will be strengthened. If only the presynaptic neuron fires with a high spike rate the synapse $W_{ij}^{CM}(t)$ is weakened after each presynaptic spike.

Plasticity of synapse $W_{ij}^{CM}(t)$ connecting neuron j of layer C to neuron i of layer M is governed by the equation

$$\frac{dW_{ij}^{CM}}{dt} = -\frac{W_{ij}^{CM}(t) - W^{CM}}{T_W} + \Omega G_j \left(\frac{\Delta t_j}{T_M} \right) \kappa(G_i, G_j) \delta_{F_i F_j}, \quad (8)$$

if $W_{ij}^{CM}(t) \in [0, 1]$. The first term describes relaxation of $W_{ij}^{CM}(t)$ to the initial value $W^{CM} = 0.2$, the respective time constant being $T_W = 1-2$ s. The second term describes changes of the synaptic strengths between neurons of the same feature type ($F_i = F_j$); the corresponding time scale is $\Omega = 0.3s^{-1}$. κ takes the value 1 if both neurons are firing within $T_M = 4$ ms and the value -1, if only the presynaptic neurons is active; in all other cases κ vanishes (for details see [5]). In case $W_{ij}^{CM}(t) \notin [0, 1]$, $W_{ij}^{CM}(t)$ relaxes to W^{CM} .

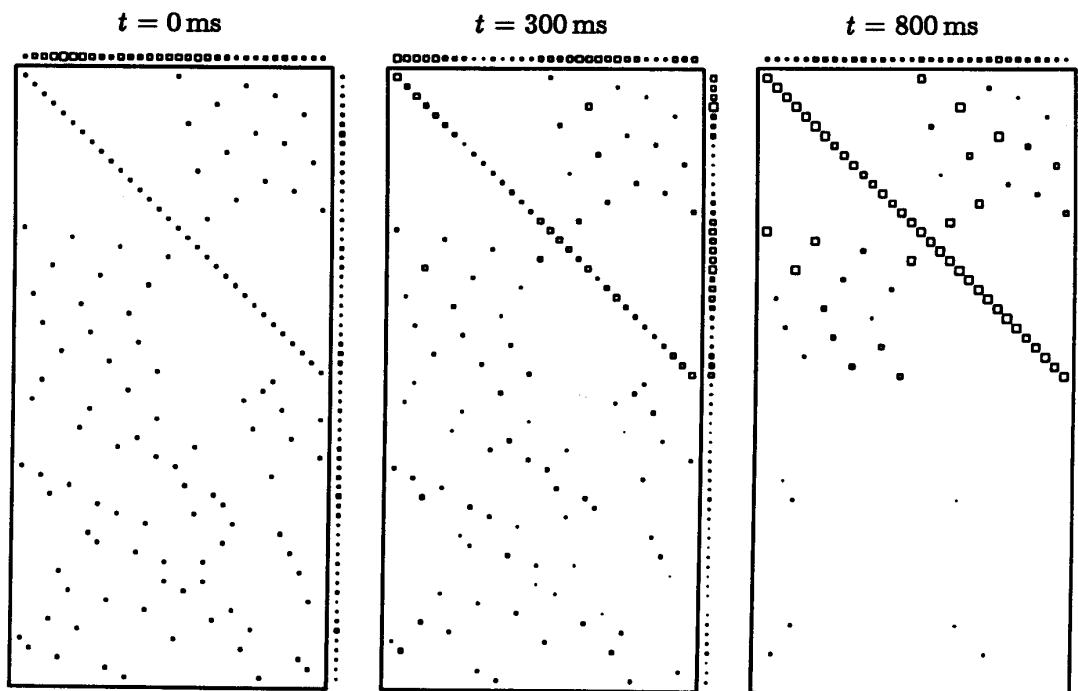


Fig. 5.
Synaptic strengths between C-layer and M-layer at time $t = 0$ ms, $t = 300$ ms, $t = 800$ ms.

6. Simulation of the system

In this section we describe a simulation of the suggested neural model. Two patterns of 32 features drawn out of 16 feature classes are stored in the memory layer. Each feature occurs twice in each pattern. The two patterns can only be distinguished according to the topology of features, i.e. according to the proper sequence of features. The pattern stored are

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 1 6 11 16 4 9 14 2 7 12 3 8 13 5 10 15 and
3 12 2 8 15 7 13 11 15 1 3 6 4 9 11 14 8 9 13 16 5 16 6 10 14 4 7 1 2 5 10 12.

In order to test the performance of the system the first pattern is presented to the system in R, respective C. The system should respond by activating only the first pattern in M, not the second pattern.

An important question for encoding feature neighborhood relations by spike correlations concerns the number f of feature classes. How many feature classes must be chosen, that a short part of the pattern is unambiguously characterized by the corresponding features? For an average cluster size b of activated neurons it is necessary to choose $f \gg b$ feature classes. Otherwise, each activated cluster is equivalent to all other clusters and the activation of layer M is rather diffuse. We cannot expect that an ordered projection from layer C to layer M arises if $f \leq b$.

In the simulation we choose the parameters ω^C , ω^M in such a way that the mean cluster size of active neurons is about $2L_e + 1$ with $L_e = 2$, $L_i = 4$. The simulation lasts from $t = -1000$ ms to $t = 1000$ ms; for $t \geq 0$ the synapses between layer C and layer M

are plastic. The mean activity of the two neural assemblies of layer **M** representing the presented pattern (bold curve) and the second pattern (thin curve) are shown in Fig. 4. For $t < 0$ the assembly representing the presented pattern fires with a slightly higher activity than the alternative assembly. After the synapses W_{ij}^{CM} became plastic the mean activity of the pattern with the correct feature topology is enhanced and activity of other neurons is suppressed. The correct pattern is selected and recognized by the system solely on account of topological correspondence.

The evolution of synapses W_{ij}^{CM} is shown at three different time instances in Fig. 5. The size of the symbols \square is proportional to the synaptic strengths. The row index marks neuron j of layer **C**, the line index marks neuron i of layer **M**. The symbols \square outside the three rectangles show the mean activity of the corresponding neurons in **C** and in **M**. At time $t = 0$ we obtain a sparse matrix of synaptic connections W_{ij}^{CM} with a non-vanishing diagonal in the upper part. During the recognition task the diagonal, which corresponds to the correct one-to-one projection between the presented and the stored pattern, is strengthened and the other synapses are reduced.

7. Summary

We have demonstrated that invariant pattern recognition of labeled patterns is possible with a two layer neural system endowed with plastic connection between the two layers. The neighborhood relations between features are encoded in spike correlations of neurons which represent these features. Thereby, patterns consisting of an equal number of features with different arrangement of these features can be distinguished. The dynamics of synapses converge fast enough that this mechanism can support invariant pattern recognition also under biological constraints.

Acknowledgement: It is a pleasure to acknowledge many and most helpful discussions with C. v.d. Malsburg on the subject of this paper.

- [1] J. J. Hopfield, Proc. Natl. Acad. Sci. USA **79**, 2554 (1982), and **81**, 3088 (1984).
- [2] E. Bienenstock and C. v.d. Malsburg, Europhys. Lett. **4**, 121; **3**, 1243 (1987).
- [3] Malsburg Cvd (1981) Internal Report 81/2, Dept. Neurobiol, MPI f. Biophysikalische Chemie, Göttingen.
- [4] D. O. Hebb *Organisation of Behavior*, (Wiley 1949).
- [5] J. Buhmann and K. Schulten, Biol. Cybern. **54**, 319 (1986); **56**, 313 (1987).