Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology)
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

A biologically inspired model for pattern recognition*

Eduardo GONZALEZ[1], Hans LILJENSTRÖM[2], Yusely RUIZ[1], Guang LI†‡[3]

([1]*Department of Biomedical Engineering, Zhejiang University, Hangzhou 310027, China*)

([2]*Department of Energy and Technology, Swedish University of Agricultural Sciences, Box 7013, SE-750, Sweden*)

([3]*National Laboratory of Industrial Control Technology, Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou 310027, China*)

†E-mail: guangli@zju.edu.cn

Received July 12, 2009; Revision accepted Nov. 4, 2009; Crosschecked Dec. 28, 2009

Abstract: In this paper, a novel bionic model and its performance in pattern recognition are presented and discussed. The model is constructed from a bulb model and a three-layered cortical model, mimicking the main features of the olfactory system. The olfactory bulb and cortex models are connected by feedforward and feedback fibers with distributed delays. The Breast Cancer Wisconsin dataset consisting of data from 683 patients divided into benign and malignant classes is used to demonstrate the capacity of the model to learn and recognize patterns, even when these are deformed versions of the originally learned patterns. The performance of the novel model was compared with three artificial neural networks (ANNs), a back-propagation network, a support vector machine classifier, and a radial basis function classifier. All the ANNs and the olfactory bionic model were tested in a benchmark study of a standard dataset. Experimental results show that the bionic olfactory system model can learn and classify patterns based on a small training set and a few learning trials to reflect biological intelligence to some extent.

Key words: Olfactory system, Neural network, Bionic model, Pattern recognition
doi:10.1631/jzus.B0910427 **Document code:** A **CLC number:** Q81

1 Introduction

Human beings have always tried to understand the world and themselves. Through the senses, touch, sight, hearing, smell, and taste, animals and particularly humans obtain information about the environment, which is used to make models of the world and cope with it in an efficient way. Scientifically, we can study the different pathways and neural substrate, the structure and dynamics of different sensory systems, in order to understand how the sensory information is processed by our brains.

One of the oldest sensory systems is the olfactory system, which is relatively simple and comparatively well known functionally and morphologically, and is thus an interesting system for understanding the process of recognition, classification, learning and recall. Many different aspects of olfaction, such as the nature of the stimuli or the mechanisms of reception and central processing, have been studied and modeled fairly extensively (Freeman, 1975; Shepherd, 1979; Freeman and Skarda, 1985; Wilson and Bower, 1992; Liljenström and Hasselmo, 1995; Aronsson and Liljenström, 2001; Haberly, 2001). However, it remains unclear how the olfactory information is recalled by the brain.

Mammalian odor information is processed by three main parts: the olfactory epithelium, the olfactory bulb and the olfactory cortex. The odorous molecules are detected by the odor receptors (ORs) and transduced into neural activation. The ORs send their axons to small structures called glomeruli within the olfactory bulb. Also, the ORs, that express common odor

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (Nos. 60874098 and 60911130129), the High-Tech Research and Development Program (863) of China (No. 2007AA042103), the National Creative Research Groups Science Foundation of China (No. 60721062) and the Project of Introducing Talents for Chinese University Disciplinal Innovation (111 Project, No. B07031)

information, converge onto one or two common glomeruli (Mombaerts *et al*., 1996; Jefferis *et al*., 2001).

Mitral cells in the olfactory bulb form synapses with the axons within glomeruli and send the odor information to the olfactory primary cortex. Mitral cells also project to granular cells inside the bulb and these receive inhibition and excitation of the mitral cells through pathways from centrifugal fibers and the olfactory cortex. When the spatiotemporal representation of an odor enters the cortex, the odor is normally identified and recalled, if it has been learned before.

Even though the olfactory neurodynamics is fairly well known, the final odor recognition process is still unclear. As one of the oldest sensory modalities in the sensory systems of mammals, olfactory nervous systems have attracted many researchers during the last decades. Some models have been developed to simulate the function of olfactory nervous systems (Freeman, 1987; Freeman and Barrie, 1994; Kay *et al*., 1995). They are useful in understanding the signal processing mechanism of olfactory systems. Therefore many mathematical models of olfactory systems have been applied to pattern recognition, often with remarkable results, which help to understand the olfactory information processing (Liljenström and Wu, 1995; Kozma *et al*., 2003; Gonzalez *et al*., 2007; Li G. *et al*., 2007; Li X. *et al*., 2006; Ma and Krings, 2009; Wu *et al*., 2009).

In this paper, a mathematical bionic model of the mammalian olfactory system is presented and its capacity for pattern recognition is tested. The full model is constructed from a bulbar model [originally developed by Li and Hopfield (1989)] which emulates the excitatory mitral and inhibitory granule cells, and from a three-layered cortical model, which mimics the structure and behavior of the piriform cortex [originally developed by Liljenström (1991)]. The main objective with this mathematical model is to mimic the principal features of the olfactory system, relating to the ability to learn and recognize input patterns, even when these are fragmented or distorted versions of the original patterns. In order to test the model capacity in classification task, a standard data set was used. The Breast Cancer Wisconsin (BCW) dataset from the University of California, Irvine (UCI) repository has been studied by other researchers in the last decades, making the results easy to compare and reproduce (Taha and Ghosh, 1997; Brunzell and Eriksson, 2000). In the following sections a description of the model and simulation results for the cancer data classification are shown.

2 Model description

Our modelling approach balances between the wish for realism when comparing theoretical and computational results with experiment, and the need for abstraction and simplification of the biological complexity for a mathematical analysis and computer simulation, in order to gain understanding of the complex structure and dynamics of the biological system.

The bionic model is based on the olfactory system and it emulates the main structural features of the system in order to create a pattern classifier. Fig. 1 shows the two main parts of the bionic model and how they are connected via feedforward and feedback

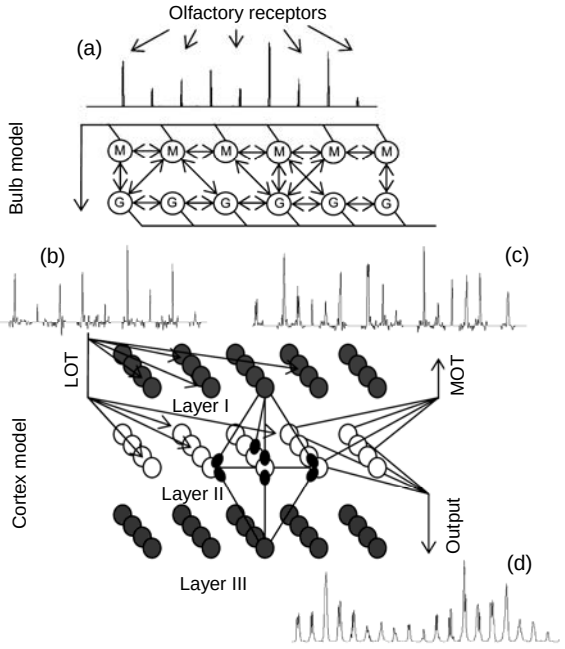

Fig. 1 The structure of our bionic model. For the bulb model during the active stage the sensory input (a) is modulated by excitatory mitral cells (M) and inhibitory granule cells (G). Olfactory information (b) passes from the olfactory bulb directly into top (Layer I) and middle (Layer II) layers on the cortex model via the lateral olfactory track (LOT). Feedback information (c) is sent up from middle cortical layer to bulbar granule nodes via medial olfactory track (MOT). The bionic model output (d) is taken from the cortical middle layer

channels. In order to give a detailed description of the bionic model, its main parts are named: bulb model and cortex model, but this does not mean that the model is an exact representation of the olfactory system.

2.1 Network architecture

In our bionic model the channel numbers of the lateral and medial olfactory tracks are equal, based on the bidirectional characteristics of the mammalian neural systems. The output of the olfactory bulb model is transmitted to the olfactory cortex model through feedforward channel connections via the lateral olfactory tract (LOT) and the feedback transmission from the cortex to glomerulus' ensembles in the bulb is sent via the medial olfactory track (MOT).

2.1.1 Bulb model

In the bulb model, based on previous works (Li and Hopfield, 1989; Li, 1990), we included N excitatory mitral cells and M inhibitory granule cells, with receptor inputs connecting directly onto the mitral cells. We disregarded the glomerular layer structure in this model. The "spherical" structure of the bulb is simplified by placing the cells in a one-dimensional ring with altogether 400 cells of each type, i.e., N=M=400.

Each cell is specified by an index, for example, the ith mitral cell and the jth granule cell for all i, j which represent the spatial locations of the cells. The ith mitral cell is the neighbor of the (i±1)th mitral cell and the (i×M/N)th granule cell. Due to the circular structure of the model, the 1st and the mth (gth) mitral (granule) cells are situated next to each other.

The synaptic connection from granule to mitral cells is represented by an N×M matrix $\boldsymbol{H}_0$, and the synaptic connection from mitral to granule cells is represented by an M×N matrix $\boldsymbol{W}_0$. For example, H_{0ij} is the connection strength from the jth granule cell to the ith mitral cell. Because we only assume local synaptic connections, $H_{0ij}\neq 0$ only if the jth granule cell and the ith mitral cell are close to each other, implying a near diagonal matrix for $\boldsymbol{H}_0$ with most non-zero elements near the diagonal, and similarly for $\boldsymbol{W}_0$. The two matrices, $\boldsymbol{H}_0$ and $\boldsymbol{W}_0$, have the same dimensions (400×400), as well as the same distribution, in our computer simulations.

The ring geometry of our bulb model implies that most of the non-zero elements are close to the matrix diagonal and corners, and we use the same distribution of the non-zero elements in $\boldsymbol{W}_0$ and $\boldsymbol{H}_0$. After creating the two matrices they are reduced to 20% ($\boldsymbol{H}_0$=$\boldsymbol{H}_0$/5 and $\boldsymbol{W}_0$=$\boldsymbol{W}_0$/5). Conversely, an N×N matrix $\boldsymbol{Z}_0$ is used to describe intralayer connection inside the mitral layer. Thus in the beginning, all mitral nodes have the same strength level of interconnection, but during the learning process these are increased (strengthened) or reduced (restricted).

2.1.2 Cortex model

The piriform cortex has a laminar structure that can be subdivided into three layers, consisting of cell bodies and connection fibers of various types. In our cortex model based on the model by Liljenström (1991), three populations of network nodes, excitatory nodes corresponding to pyramidal cells and two kinds of inhibitory nodes corresponding to interneurons are likewise organized into three layers. The most superficial layer of the cortex, Layer I, contains inhibitory interneurons and afferent fibers from the olfactory bulb, making excitatory connections with these nodes.

The interneurons in Layer I are usually referred to as feedforward inhibitory cells (Shepherd, 1979; Wilson and Bower, 1989), although they also likely to receive feedback connections from nearby pyramidal cells (Hasselmo *et al*., 1990) from Layer II. Cell bodies of excitatory pyramidal cells are the main constituents of the middle layer, Layer II. The deepest layer of the olfactory cortex contains both pyramidal cells and interneurons, but to simplify the cortical model, only feedback inhibitory nodes are included on Layer III.

Pyramidal cells make excitatory connections with interneurons in Layer I and Layer III. These neurons or ensemble of neurons make feedforward and feedback inhibitory connections with the pyramidal cells of Layer II (Fig. 2a), respectively.

In the cortical model there are N_{I} feedforward inhibitory nodes in the upper cortical layer, N_{II} excitatory nodes in the middle cortical layer, and N_{III} feedback inhibitory nodes in the bottom cortical layer, with N_{layer}=N_{I}=N_{II}=N_{III} and N_{C}=3×N_{layer}. In our simulations N_{layer}=400 nodes (for 20×20 layer dimensions) and N_{C}=1200 nodes. There are interlayer connections between Layer I and Layer II and

between Layer II and Layer III. Intralayer connections exist only within Layer II, where an extensive association fiber system connects excitatory nodes with each other, both towards the back of the network, caudally, and towards the front, rostrally (Fig. 2b). The fibers spread out radially from the originating cells. Inhibitory-inhibitory connections may exist in the real cortex but have been excluded in this work.

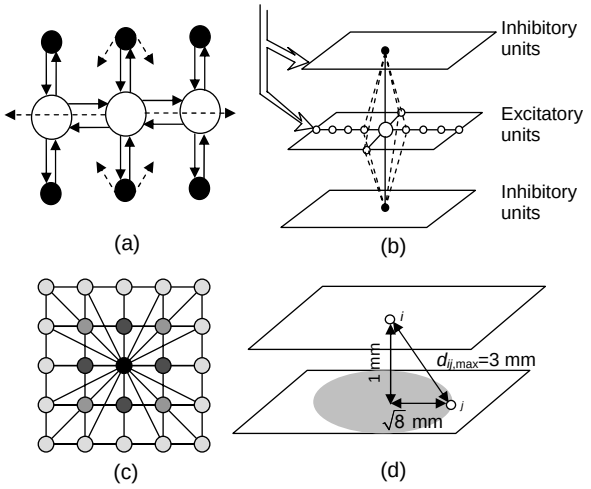

Fig. 2 Architecture of the cortical model. (a) General principle for connections between excitatory nodes (open circles) and inhibitory nodes (filled circles). Dashed lines indicate that connection strength decrease with distance. The model is taken from Liljenström (1991); (b) Overall structure of the model olfactory cortex with intralayer and interlayer connections. The arrows represent the input from the bulb model. The model is taken from Liljenström (1991); (c) The excitatory nodes in middle layer, which make connections with inhibitory nodes in upper or bottom layer; (d) Maximum range ($d_{ij,max}$) of 3 mm for inhibitory-excitatory connections

The synaptic strength between any two nodes j and i is given by a connection weight w_{ij}. In order to model a presumed decrease in a cell's influence with distance, the connection strength is attenuated exponentially over distance. For excitatory-excitatory connections the strength asymptotically approaches a minimum value. Inhibitory nodes in the top layer only make local connections with excitatory nodes in middle layer within a 3-mm range, corresponding to 25 adjacent nodes (Fig. 2c). Also the connection probabilities are included in the model expression of the connection weights w_{ij}:

$$w_{ij}=p_{ij}c_S\exp[(-d_{ij}/\lambda_S+a_S)(1+a_S)^{-1}], \quad (1)$$

where p_{ij} is the probability that a connection is made between nodes i and j and c_S is a constant determining the synaptic strength in the connections of type S between nodes in Layer II or between nodes in Layer II and in Layer I or III. d_{ij} is the radial distance, measured in millimeters, between nodes i and j (Fig. 2d) and has to be lower than the maximum connection range (in our simulation $d_{ij}\leq 3$ mm). λ_S is a space constant and a_S is a value which determines the minimum connection strength that is approached asymptotically ($a_S\neq 0$ only for excitatory-excitatory connections).

For our simulation the w_{ij} equation is simplified to:

$$w_{ij}=\begin{cases} CL_{hk}\exp(-d_{ij}/\lambda_S), & P_{hk}<T_{hk} \\ 0, & P_{hk}\geq T_{hk} \end{cases}, \quad (2)$$

where CL_{hk} is a constant determining the synaptic strength in the connections between nodes in Layer II or between nodes in Layer II and in Layer I or III. P_{hk} is the random value and T_{hk} is the probability connection threshold, below which the synaptic connection between node i in Layer h and node j in Layer k is made.

Olfactory information reaches the olfactory cortex from the olfactory bulb via a bundle of nerves, LOT. These fibers spread out and make excitatory connections with feedforward inhibitory cells and pyramidal cells on Layer I and Layer II. The output of the bulb model is modelled in our work with N_{LOT}= 400 separated input connections.

Connection weights and parameters used in the model are only loosely constrained by physiological data and are chosen freely in order to reach the best performance in the pattern classification tasks.

2.2 Network dynamics

2.2.1 Bulb model

The olfactory bulb can be viewed as the first central olfactory relay station extracting specific stimulus features, a function characteristic of the primary sensory areas in the brain (Doty, 2003; Lowe, 2003; de Araujo *et al*., 2005). The cellular structure of the bulb is well established and in this work, the olfactory bulb was modelled using a simple approximation of excitatory mitral and inhibitory granule cells. The activity of mitral cells was spatially distributed such that odorants were represented in the

bulb model by a distributed pattern of mitral cell activity (Mori *et al*., 1999; Leon and Johnson, 2003). Mitral cells adjacent to each other project to the same or neighboring glomerulus.

The dynamics of each neural ensemble is described using a first order differential equation based on physiological experiments of the olfactory system (Freeman, 1975). The mitral and granule layers on the model are mathematically represented by ordinary differential equations (ODEs).

For mitral layer:

$$\mathrm{d}m_i/\mathrm{d}t=-am_i+I_{\mathrm{w}}(t)+I_{\mathrm{e}i}(t)+r, \quad (3)$$

$$I_{\mathrm{w}}(t)=\sum_{j=1}^{N}H_{0ij}Q_j[g_i(t-\delta_{ij})]+\sum_{j=1}^{N}M_{0ij}Q_j[m_j(t-\delta_{ij})]; \quad (4)$$

For granule layer:

$$\mathrm{d}g_i/\mathrm{d}t=-bg_i+\sum_{j=1}^{N}W_{0ij}Q_j[m_j(t-\delta_{ij})]+cI_{\mathrm{c}i}(t)+r. \quad (5)$$

In Eqs. (3)–(5):

$$Q_j(x)=S_x\tanh[(x-Th)/S_x],$$
$$\text{if } x\le Th, \text{ then } S_x=10S_x; \text{ if } x>Th, \text{ then } S_x=7S_x. \quad (6)$$

a, b and c represent rate constants. The external input to the mitral layer is symbolized by I_{e} and I_{c} represents the cortical feedback input to granule layer. $m_i(t)$ and $g_j(t)$ symbolize the dynamic state of the ith and jth neural ensemble in the mitral and granule layer, respectively. The positive, negative or zero value of the connection strength, M_{0ij}, H_{0ij}, or W_{0ij}, represents the weight connection into the mitral layer and between the mitral and granule layers from neural population j to i; therefore, these matrix values define the system topology. $Q(x)$ is the asymmetric sigmoidal input/output transformation function used for mitral and granule cells, which was derived from the Hodgkin-Huxley equations, Th is a threshold defining the sigmoid shape, and S_x represents the maximum asymptote of the sigmoid function (Freeman, 1979; Freeman and Skarda, 1985). δ corresponds to the specific axonal transmission delay and r represents the noise or spontaneous neural activity.

The input-output relations [Eq. (6)] for the mitral and granule cells, respectively, are experimentally derived by Freeman (1979) and plotted in Fig. 3, with the threshold Th=1. However, the exact form of these relations is not essential to the system behavior, as long as the shape is qualitatively preserved. Since granule cells do not have axons, they are modelled using a larger linear range, and thus a less strong nonlinear threshold effect, than mitral cells (Fig. 3). In our simulation S_x=0.14 for granule nodes and S_x=0.29 for mitral nodes.

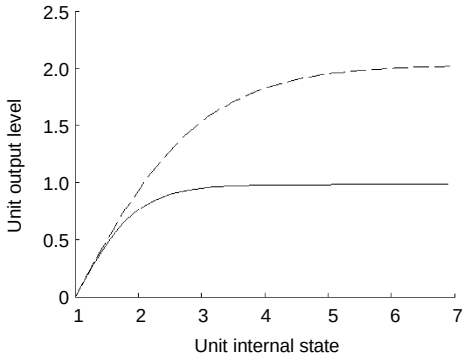

Fig. 3 Non-linear input-output functions for mitral (solid line) and granule (dashed line) nodes, based on Freeman and Skarda (1985)

2.2.2 Cortex model

Basically the piriform cortex structures belonging to allocortex are thinner and structurally less complex (having three cortical layers) than the neocortex (Liljenström, 1991; Wilson and Bower, 1992; Kowianski *et al*., 1999; Johnson *et al*., 2000). The architecture of the cortical model was based on the three-layered structure of the olfactory piriform cortex (Liljenström, 1991), using similar network connectivity, but relatively simple model nodes, representing populations of neurons.

The two sets of inhibitory nodes have two different time constants and slightly different connectivity to the excitatory nodes. All connections were modelled with distance-dependent time delays for signal propagation, corresponding to the geometry and fiber characteristics of the real cortex.

The dynamics of a network node is denoted by the mean membrane potential u and is given by a set of coupled nonlinear first-order differential delay equations:

$$\mathrm{d}u_i/\mathrm{d}t=-u_i/\tau_i+\sum_{j\neq i}^{N}w_{ij}g_j[u_j(t-\delta_{ij})]+I_i(t)+r, \quad (7)$$

where τ characterizes the spontaneous decay, w_{ij} is the connection weight and δ corresponds to the specific

axonal transmission delay. The external input to the upper and middle cortical layers from the mitral layer is represented by I and r is noise or spontaneous neural activity. The input-output relationship of a neuron population in the piriform cortex, $g_j(u_j)$, is a continuous sigmoid function determined by recording evoked potentials (Freeman, 1975):

$$g_j(u_j)=CQ_j\left\{1-\exp\left[\frac{-\exp(u_j)-1}{Q_j}\right]\right\}, \qquad (8)$$

where the gain parameter Q_j is assumed to correspond to the level of any particular neuromodulator, such as acetylcholine. C is a normalization constant (in our computer simulation C=0.05). In this case, neural adaptation was implemented as an exponential decay of the output proportional to the time average of previous output, and thus the input-output relationship becomes:

$$g_{j\alpha}=g_j(u_j)\exp\left\{-\left[\alpha\left\langle g_j(u_j)\right\rangle_T\right]^2\right\}, \qquad (9)$$

where $\langle\ \rangle_T$ denotes the time average over the last T ms, and α is an adaptation parameter (Liljenström and Hasselmo, 1995; Liljenström, 2003).

2.3 Network learning

While time constants, signal velocities, and other system parameters are determined by physiological constraints, the connection weights should be adjusted properly for the best performance of the model. The same kind of model behavior can be found within several parameter regimes, but when the choice of values is limited by biological constraints, the variation in connection strengths also becomes more restricted.

For our bionic model, two learning processes are used: Hebbian associative learning and habituation. These learning processes exist in a subtle balance and their relative importance changes at various stages of the memory process. The memory basins and attractors are formed via Hebbian learning under reinforcement, while the impact of environment noise including the background inputs without any information is reduced by habituation. The learning processes are applied to the bulb model and the cortex model.

In the bionic model the learning rules are implemented into the following formulae. The output of the bulb model is used to measure the activity at the mitral layer (M) and the middle layer, Layer II (L), of the cortical model is taken as the activity measure. Over a period the presentation of input patterns (here the period of the input presentation is 200 ms), the mean standard deviation of the output of the ith node SD_{ai} is used to present the activity of the ith channel (M_i and L_i for bulb and cortical models, respectively).

$$SD_{ai}=1/S\sum_{k=1}^{S}SD_k\ . \qquad (10)$$

The response output period is divided into S equal segments. Accordingly, SD_k is the standard deviation of the kth segment, and SD_{ai} is the mean value of all these S segments. $\boldsymbol{SD}_a$ is an $(1\times N)$ vector representing the activity of all nodes in the mitral or middle cortical layer (in our computer simulation N=400). Thus $\langle\boldsymbol{SD}_a\rangle$ is the mean activity over the period of the presentation of input patterns:

$$\langle\boldsymbol{SD}_a\rangle=1/N\sum_{i=1}^{N}SD_{ai}\ . \qquad (11)$$

According to modified Hebbian rule, each pair of M or L nodes co-activated by the stimulus has their lateral connections strengthened, where $w(m)_{ij}$ or $w(L)_{ij}$ represents the connection within mitral or cortical middle layer respectively and $w(m\text{-}m)_{ij}$ or $w(II\text{-}II)_{ij}$ indicates the weights of the connections from M_i to M_j and from M_j to M_i or from L_i to L_j and from L_j to L_i. The nodes with activities larger than the mean on the layer are considered as activated ones and strengthened with the Hebbian coefficient (H_{heb}). In contrast those with activity less than the mean are not considered to be activated ones and these connections are decreased by the Habituation coefficient (H_{hab}) and the simulation period (T). A bias coefficient K is defined in the modified Hebbian learning rule to avoid the weight space saturation (Wang *et al*., 2005):

Begin

If $\left(SD_{ai}>(1+K)\langle\boldsymbol{SD}_a\rangle;SD_{aj}>(1+K)\langle\boldsymbol{SD}_a\rangle;i\neq j\right)$

$w(m)_{ij}=H_{heb}w(m)_{ij}$;

Else if $\left(SD_{ai} > (1+K)\left\langle \boldsymbol{SD}_{a}\right\rangle; SD_{aj} > (1+K)\left\langle \boldsymbol{SD}_{a}\right\rangle; i = j\right)$

$w(m)_{ij}$=0;

Else $w(m)_{ij}=H_{\text{hab}}\text{^}(Tw(m)_{ij})$.

End

The same procedure is used for the bulbar mitral layer and middle cortical layer. For the habituation learning, a decrease in excitatory output synapses occurs continuously at each time of stimulus presentation for every node that receives input without reinforcement. Unlike the Hebbian learning, the reduction is reversible and does not require pairwise activation.

Several samples for each class are given during training. At the end of the training session the connection weights are fixed for the classification test. The amplitude of N activity outputs taken from the middle layer of the cortical model is calculated and defined as a feature vector for every trial, as well as the mean activity of those trials belonging to the same class.

The feature vector corresponds to a point in N-space. A cluster of points in N-space is formed by a set of training trials within the same stimulus class. The mean is defined as the gravity center of the cluster representing the class. The inputs of different classes form multiple clusters of feature vectors, each with its own gravity center. When a test pattern is given, its feature vector is calculated and the Euclidean distances from the corresponding feature vector to those trained pattern cluster centers are calculated. The class of the pattern is determined by the minimum distance to a center (Wang *et al*., 2005).

3 Results

In order to demonstrate the pattern classification performance of the bionic model, a standard test set, BCW, available from the UCI Machine Learning Data Repository (Wolberg, 1991) was selected. These data were gathered by William H. Wolberg from the University of Wisconsin Hospital, and each sample consists of nine different features of breast tumors with values ranging from 1 to 10. The cancer data were grouped into two classes, benign (444 samples) and malignant (239 samples) for a total of 683 complete data. The main reason for choosing this specific dataset is that it is available for public download and it has been studied by other researchers, making the results easily comparable (Brunzell and Eriksson, 2000).

Nine feature channels were used as input into the mitral layer in the bulb model. The bottom layer of the cortical model provided inhibitory feedback to the middle excitatory layer, and the top layer sent feedback signals to the granule cells from the bulb model. Cortical model output was taken from the middle layer, and was classified using a Euclidean distance algorithm. After the training process, the centers of gravity for two classes (benign and malign) were calculated and the connection weights were fixed. The minimum Euclidean distance between the centers for each class during the training process and the cortical output during the test process were used as a decision criterion for the final recognition of each kind of cancer. Furthermore, during the classification stage, the connection weights of the cortical model were fixed to the values obtained during the training stage.

The performance of the bionic model was compared to three standard artificial neural networks (ANNs): a back-propagation network (BPN), a support vector machine (SVM) classifier, and a radial basis function (RBF) classifier. The ANNs as well as the bionic model took part in the benchmark study of a standard dataset.

Before classifying the BCW dataset with the novel bionic model, the cortical surface that receives input from the bulb model and the size of training set were optimized. Also some features of the ANNs were optimized in order to minimize the number of misclassifications during the tests.

3.1 Optimization of the feedforward channels

We optimized the bionic model with respect to the number of inputs to the cortical surface from the olfactory bulb via the feedforward channels. It has been reported that in monkeys the olfactory input to the entorhinal cortex is about 12.5% of that to the entorhinal surface area (Amaral *et al*., 1987), whereas in rats it is at least half (Haberly, 2001).

Initially, the bulb and cortical models were connected by a full connection. In order to find the optimal number of feedforward channels arriving with signal information from the bulb to the cortex, the number of channels was varied from 9 to 400 as SN^2,

where *SN*=(3, 4, 5, …, 20). Furthermore, the channels were placed into the cortical layers equidistantly, as shown in Fig. 4. Keeping the rest of the model parameters constant, the area of the cortical surface that received input from the bulb was varied from a minimum (2.5%) to a maximum area (100%), where each configuration was tested fifteen times. The correct classification rate of the model was used as an indicator of the optimal area amount.

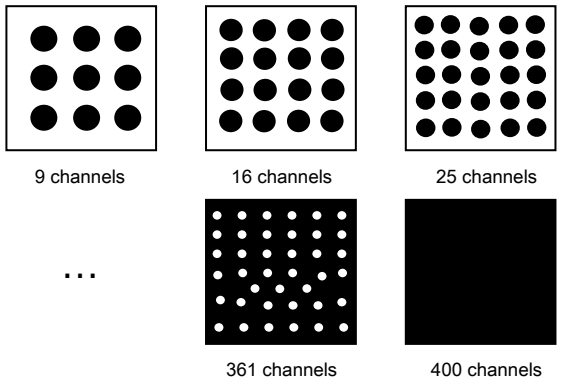

Fig. 4 Cortical areas (black) receiving input from olfactory bulb

In order to optimize the configuration of the bionic model, the classification rates with different configurations of connection were calculated for statistical testing. As the classification rates did not follow a normal distribution, the non-parametric method, Kruskal-Wallis test (Kruskal and Wallis, 1952), was applied to test the statistical difference when the cortical surface receiving input from the Mitral layer of the bulb model varied. The statistical significance of the differences in the classification rates across the fifteen observations between the groups shaped by the configuration of cortical surface receiving input from bulb model was evaluated. Over the eighteen groups the *P* value was lower than 0.01, demonstrating that at least one group was significantly different from the others. The Kruskal-Wallis test also revealed a significant difference between the configurations for numbers of channels higher and lower than 35% of the cortical surface. Therefore, based on Kruskal-Wallis test results, any configuration higher than 35% will produce better classification rates. Due to the clear peak that appeared at the cortical surface optimization (Fig. 5), the 38% of cortical surface receiving input was selected as the optimal configuration.

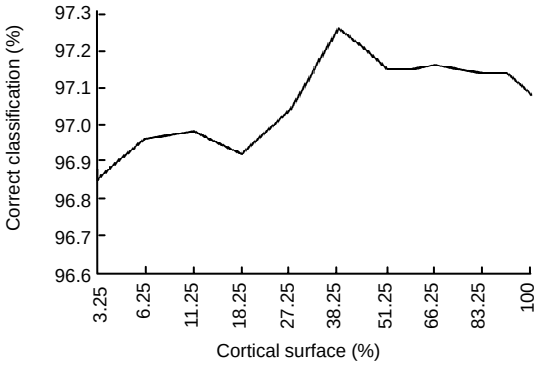

Fig. 5 Optimization of bionic model channels for BCW dataset

3.2 Creation of the ANNs

The first ANN used in our benchmark was a BPN, which has been one of the most studied and used algorithms for neural networks learning. It minimizes the error function in weight space using the method of gradient descent. Its structure is relatively simple and highly related to the application under consideration. In our experiments we created a BPN with three-layered (input-hidden-output) structure, and usually one hidden layer suffices for many applications (Patterson, 1996; Duda *et al*., 2001). The number of nodes for the input layer was set at 9 nodes (BCW dataset dimension) and the number of nodes for the output layer was fixed at 2 nodes (kind of cancers, benign and malign). As BPNs are sensitive to the number of neurons in their hidden layers (too few neurons can lead to underfitting and too many neurons can contribute to overfitting) and due to the fact that there is no explicit rule to determine the number of hidden nodes, a test varying the hidden neuron number was undertaken and the best result was used for the configuration of twelve hidden nodes.

The second ANN tested was an SVM classifier, which is a novel type of learning machine based on statistical learning theory (Cortes and Vapnik, 1995). It has gained popularity during the last decade due to its many attractive features and promising empirical performance. The formulation embodies the Structural Risk Minimisation principle, as opposed to the Empirical Risk Minimisation approach commonly employed within statistical learning methods. Given a set of points that belong to either of two classes, an SVM finds the hyperplane leaving the largest possible fraction of points of the same class on the same side,

while maximizing the distance of either class from the hyperplane. In our study the SVM classifier with a linear kernel function was designed.

The last ANN used in our work was the RBF classifier. These classifiers emerged as a variant of ANNs in the late 80's (Margalit, 1988; Renals, 1989), but their roots are entrenched in much older pattern recognition techniques. The RBF has an associated width parameter, which is related to the spread of the function around its center. The optimal value of the spread parameter must be found for each given problem. For this reason, and since it is related to the specific data to be classified, the RBF classifier was optimized with respect to the spread parameter. Varying the spread parameter, the BCW data were classified, and the best classification results were found for a spread value of 30 (Fig. 6).

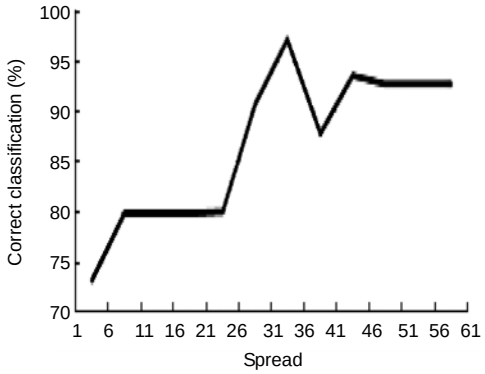

Fig. 6 Optimization of spread parameter for the RBF

3.3 Optimization of training set

Other optimization tasks were required to determine the optimal training set size, showing how fast the novel bionic model and the other three ANNs could learn and recognize the two types of cancer patterns (benign and malign). Picking the minimum size of the training set for an ANNs is a challenge. A training set that is too large leads to long training time and a too small set results in unstable learning.

For the cancer data classification, a training set with 110 samples per class was created and the number of samples used during the training of each network was varied in order to find the optimal number of training samples. After each network had been trained, the whole BCW dataset, consisting of 683 patterns, was classified.

Fig. 7 shows the performance of the BPN (Fig. 7a), SVM classifier (Fig. 7b), RBF classifier (Fig. 7c), and the olfactory bionic model (OBM) (Fig. 7d) using training set of different sizes during the training process. The OBM exhibited the best learning capacity, reaching the optimal value (97.45%) using fewer samples (13 instances per class) during the training process when the other ANNs needed a large number of data points to achieve good classification results.

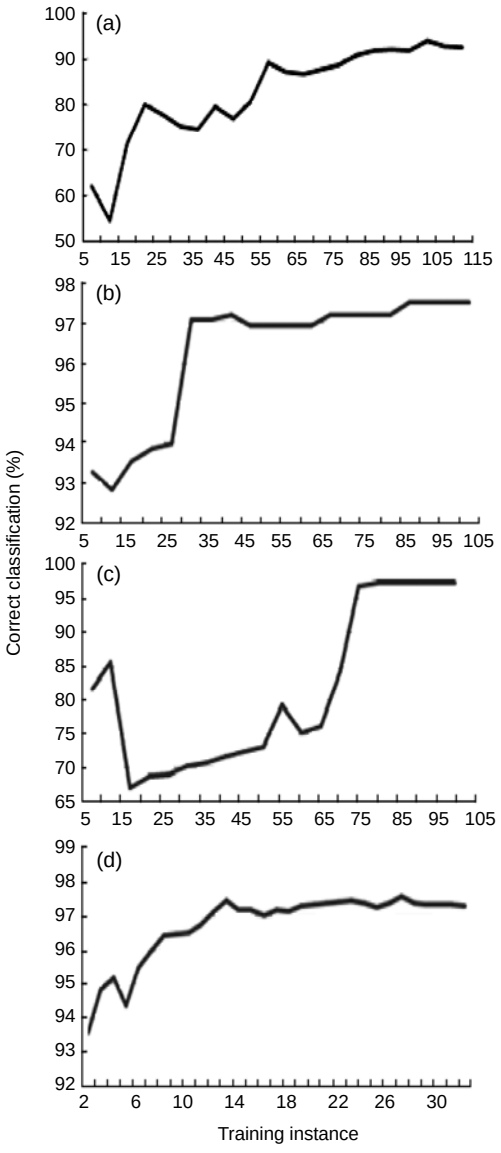

Fig. 7 Optimization of the training instances per class for BCW dataset. (a) BPN; (b) SVM classifier; (c) RBF classifier; (d) OBM

3.4 Classification rates

Table 1 shows the results of the benchmark study for the three classical ANNs and the OBM. The BPN

was the classifier with the worst results with respect to classification rate or training set size. The SVM and RBF classifiers achieved good results, but the size of the training set required in both cases (24.89% of BCW dataset) was higher than that required by our bionic model (7.62% of BCW dataset). Thus the OBM was able to learn and recognize the cancer data, achieving the best results and demonstrating its capacity in pattern recognition tasks.

Table 1 Summary of the best classification for BCW dataset

Method	Best correct classification		
	Training instances per class	Average error	Correct classification (%)
BPN	100	41.95	93.85
SVM	85	17.00	97.51
RBF	85	18.00	97.36
OBM	26	16.67	97.56

Even when the training time of all these ANNs was shorter than that of our bionic model, they used a training set volume higher than the bionic model. Our bionic model, as a form of biological intelligence, reached the best correct classification (97.56%) and a remarkable performance (classification rate higher than 97%) for a small training sets with at least 13 instances per class (3.81% of BCW dataset). Meanwhile the ANN classification rate using 15 instances on the training set (Table 2) was poorer than that of the bionic model (BPN: 71.39%, SVM: 93.55%, RBF: 67.34%, and OBM: 97.19%). The achieved results show the model excellence in a practical pattern classification task and justify the efforts to embody its ODE in analog very large scale integration (VLSI) (Principe *et al*., 2001).

Table 2 Summary of classification for BCW dataset using small training set

Method	Best correct classification		
	Training instances per class	Average error	Correct classification (%)
BPN	15	195.40	71.39
SVM	15	44.00	93.55
RBF	15	223.00	67.34
OBM	15	19.13	97.19

The OBM results were better than those achieved by the ANNs, as regards both the training set size and correct classification rate. Furthermore, the performance of the OBM sorting BCW dataset was even better than previous results reported by Brunzell and Eriksson (2000) after several feature reduction methods, like Mahalanobis based linear transformation (97.1% of correct classification), canonical variables (96.6% of correct classification), Young, Marco and Odell's method (96.5% of correct classification), and principal component analysis (96.6% of correct classification).

4 Discussion

In this work, a novel model mimicking the main features of the olfactory system has been analyzed, and its performance on the BCW dataset classification demonstrated. Our bionic model is constructed from two principal parts: a bulb model and a cortex model. The bulb model is composed of mitral and granule cells, whereas the cortical model mimics the three-layered structure of the mammalian piriform cortex.

A standard set with breast cancer data was used to test the pattern recognition capacity of our model. The performance of the bionic model was also compared with three classical ANNs: a BPN, an SVM classifier, and an RBF classifier. Computer simulation using our OBM showed its great capacity to learn and classify patterns based on a very small training set and a few learning trials, reflecting remarkable biological intelligence to an extent. However, when considering the memory requirements and the computational time to convergence, our bionic model still cannot compete against conventional ANNs for solving practical problems. The current digital computers are the bottleneck for the models based on biological neurous systems due to the time required to solve ODE by numerical integration. In this sense, the implementation of the bionic models in analog VLSI (Principe *et al*., 2001) and random graph theory arise as promising research.

References

Amaral, D.G., Insausti, R., Cowan, W.M., 1987. The entorhinal cortex of the monkey: I. Cytoarchitectonic organization. *The Journal of Comparative Neurology*, **264**(3):326-355. [doi:10.1002/cne.902640305]

Aronsson, P., Liljenström, H., 2001. Effects of non-synaptic neuronal interaction in cortex on synchronization and

learning. *Biosystems*, **63**(1-3):43-56. [doi:10.1016/S0303-2647(01)00146-0]

Brunzell, H., Eriksson, J., 2000. Feature reduction for classification of multidimensional data. *Pattern Recognition*, **33**(10):1741-1748. [doi:10.1016/S0031-3203(99)00142-9]

Cortes, C., Vapnik, V., 1995. Support-vector networks. *Machine Learning*, **20**(3):273-297. [doi:10.1007/BF00994018]

de Araujo, I.E., Rolls, E.T., Velazco, M.I., Margot, C., Cayeux, I., 2005. Cognitive modulation of olfactory processing. *Neuron*, **46**(4):671-679. [doi:10.1016/j.neuron.2005.04.021]

Doty, R.L., 2003. Handbook of Olfaction and Gustation. Dekker, New York, p.235-276. [doi:10.1016/S0165-5876(96)90071-3]

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern Recognition. John Wiley & Sons, New York, p.282-349. [doi:10.1007/s00357-007-0015-9]

Freeman, W.J., 1975. Mass Action in the Nervous System. Academic Press, New York, p.64-120.

Freeman, W.J., 1979. Nonlinear gain mediating cortical stimulus-response relations. *Biological Cybernetics*, **33**(4):237-247. [doi:10.1007/BF00337412]

Freeman, W.J., 1987. Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. *Biological Cybernetics*, **56**(2-3):139-150. [doi:10.1007/BF00317988]

Freeman, W.J., Skarda, C.A., 1985. Spatial EEG patterns, non-linear dynamics and perception: the neo-sherringtonian view. *Brain Research Reviews*, **10**(3):147-175. [doi:10.1016/0165-0173(85)90022-0]

Freeman, W.J., Barrie, J.M., 1994. Chaotic Oscillations and the Genesis of Meaning in Cerebral Cortex. *In*: Buzsaki, G., Llinas, R., Singer, W., Berthoz, A., Christen, Y. (Eds.), Temporal Coding in the Brain. Springer, Berlin, p.13-37.

Gonzalez, E., Li, G., Ruiz, Y., Zhang, J., 2007. A Tea Classification Method Based on an Olfactory System Model. Advances in Cognitive Neurodynamics ICCN 2007. Springer, the Netherlands, p.747-751. [doi:10.1007/978-1-4020-8387-7_129]

Haberly, L.B., 2001. Parallel-distributed processing in olfactory cortex: new insights from morphological and physiological analysis of neuronal circuitry. *Chemical Senses*, **26**(5):551-576. [doi:10.1093/chemse/26.5.551]

Hasselmo, M.E., Wilson, M.A., Anderson, B.P., Bower, J.M., 1990. Associative memory function in piriform (olfactory) cortex: computational modeling and neuropharmacology. *Cold Spring Harbor Symposia on Quantitative Biology*, **55**:599-610.

Jefferis, G.S., Marin, E.C., Stocker, R.F., Luo, L., 2001. Target neuron prespecification in the olfactory map of Drosophila. *Nature*, **414**(6860):204-208. [doi:10.1038/35102574]

Johnson, D.M.G., Illig, K.R., Behan, M., Haberly, L.B., 2000. New features of connectivity in piriform cortex visualized by intracellular injection of pyramidal cells suggest that "primary" olfactory cortex functions like "association" cortex in other sensory systems. *The Journal of Neuroscience*, **20**(18):6974-6982.

Kay, L., Shimoide, K., Freeman, W.J., 1995. Comparison of EEG time series from rat olfactory system with model composed of nonlinear coupled oscillators. *International Journal of Bifurcation and Chaos*, **5**(3):849-858. [doi:10.1142/S0218127495000636]

Kowianski, P., Lipowska, M., Morys, J., 1999. The piriform cortex and the endopiriform nucleus in the rat reveal generally similar pattern of connections. *Folia Morphol (Warsz)*, **58**(1):9-19.

Kozma, R., Freeman, W.J., Erdi, P., 2003. Basic principles of the KIV model and its application to the navigation problem. *Journal of Integrative Neuroscience*, **2**(1):125-140. [doi:10.1142/S0219635203000159]

Kruskal, W.H., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, **47**(260):583-621. [doi:10.2307/2280779]

Leon, M., Johnson, B.A., 2003. Olfactory coding in the mammalian olfactory bulb. *Brain Research Reviews*, **42**(1):23-32. [doi:10.1016/S0165-0173(03)00142-5]

Li, G., Jin, Z., Freeman, W.J., 2007. Mandarin Digital Speech Recognition Based on a Chaotic Neural Network and Fuzzy C-means Clustering. Fuzzy Systems Conference. FUZZ-IEEE 2007, July 23-26, 2007, p.1-5. [doi:10.1109/FUZZY.2007.4295337]

Li, X., Li, G., Wang, L., Freeman, W.J., 2006. A study on a bionic pattern classifier based on olfactory neural system. *International Journal of Bifurcation and Chaos*, **16**(8):2425-2434. [doi:10.1142/S0218127406016173]

Li, Z., 1990. A model of olfactory adaptation and sensitivity enhancement in the olfactory bulb. *Biological Cybernetics*, **62**(4):349-361. [doi:10.1007/BF00201449]

Li, Z., Hopfield, J.J., 1989. Modeling the olfactory bulb and its neural oscillatory processings. *Biological Cybernetics*, **61**(5):379-392. [doi:10.1007/BF00200803]

Liljenström, H., 1991. Modeling the dynamics of olfactory cortex using simplified network units and realistic architecture. *International Journal of Neural Systems*, **2**(1/2):1-15. [doi:10.1142/S0129065791000029]

Liljenström, H., 2003. Neural stability and flexibility: a computational approach. *Neuropsychopharmacology*, **28**(S1):S64-S73. [doi:10.1038/sj.npp.1300137]

Liljenström, H., Hasselmo, M.E., 1995. Cholinergic modulation of cortical oscillatory dynamics. *Journal of Neurophysiology*, **74**(1):288-297.

Liljenström, H., Wu, X.B., 1995. Noise-enhanced performance in a cortical associative memory model. *International Journal of Neural Systems*, **6**(1):19-29. [doi:10.1142/S0129065795000032]

Lowe, G., 2003. Electrical signaling in the olfactory bulb. *Current Opinion in Neurobiology*, **13**(4):476-481. [doi:10.1016/S0959-4388(03)00092-8]

Ma, Z., Krings, A.W., 2009. Insect sensory systems inspired computing and communications. *Ad Hoc Networks*, **7**(4):742-755. [doi:10.1016/j.adhoc.2008.03.003]

Margalit, M., 1988. Surface Fitting and Compression of Two-dimensional Scattered Data. International Conference on

Acoustics, Speech, and Signal Processing, ICASSP-88. [doi:10.1109/ICASSP.1988.196716]

Mombaerts, P., Wang, F., Dulac, C., Chao, S.K., Nemes, A., Mendelsohn, M., Edmondson, J., Axel, R., 1996. Visualizing an olfactory sensory map. *Cell*, **87**(4):675-686. [doi:10.1016/S0092-8674(00)81387-2]

Mori, K., Nagao, H., Yoshihara, Y., 1999. The olfactory bulb: coding and processing of odor molecule information. *Science*, **286**(5440):711-715. [doi:10.1126/science.286.5440.711]

Patterson, D.W., 1996. Artificial Neural Networks: Theory and Applications. Singapore, Prentice Hall.

Principe, J.C., Tavares, V.G., Harris, J.G., Freeman, W.J., 2001. Design and implementation of a biologically realistic olfactory cortex in analog VLSI. *Proceedings IEEE*, **89**(7):1030-1051. [doi:10.1109/5.939813]

Renals, S., 1989. Radial basis function network for speech pattern classification. *Electronics Letters*, **25**(7):437-439. [doi:10.1049/el:19890300]

Shepherd, G.M., 1979. The Synaptic Organization of the Brain. Oxford University Press, New York, p.377-416. [doi:10.1093/acprof:oso/9780195159561.001.1]

Taha, I., Ghosh, J., 1997. Characterization of the Wisconsin Breast Cancer Database Using a Hybrid Symbolic-Connectionist System. Proceedings of ANNIE'96.

Wang, L., Li, G., Liu, X., Wang, B., Freeman, W.J., 2005. Study of a Chaotic Olfactory Neural Network Model and Its Applications on Pattern Classification. Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Jan. 17-18, J2006, p.3640-3643. [doi:10.1109/IEMBS.2005.1617270]

Wilson, M.A., Bower, J.M., 1989. The Simulation of Large-scale Neural Networks. Methods in Neuronal Modeling: From Synapses to Networks. MIT Press, p.291-333.

Wilson, M.A., Bower, J.M., 1992. Cortical oscillations and temporal interactions in a computer simulation of piriform cortex. *Journal of Neurophysiology*, **67**(4):981-995.

Wolberg, W.H., 1991. Wisconsin Diagnostic Breast Cancer, Available from http://archive.ics.uci.edu/ml/datasets [Accessed on Aug. 6, 2008].

Wu, C., Chen, P., Yuan, Q., Wang, P., 2009. Response enhancement of olfactory sensory neurons-based biosensors for odorant detection. *J. Zhejiang Univ. Sci. B*, **10**(4):285-290. [doi:10.1631/jzus.B0820220]