ASSOCIATIVE MEMORIES FOR CHEMICAL SENSING

Reznik A.M.\textsuperscript{1}, Shirshov Yu. M.\textsuperscript{2}, Snopok B.A.\textsuperscript{2},
Nowicki D.W.\textsuperscript{1}, Dekhtyarenko A.K.\textsuperscript{1}, Kruglenko I. V.\textsuperscript{2}

\textsuperscript{1}Institute of Mathematical Machines and Systems of NAS of Ukraine
\textsuperscript{2}Institute of Semiconductor Physics of NAS of Ukraine
neuro@immsp.kiev.ua

ABSTRACT

We consider application of neural associative memories to chemical image recognition. Chemical image recognition is identification of substance using chemical sensors' data. The primary advantage of associative memories as compared with feed-forward neural networks is high-speed learning. We have made experiments on odour recognition using hetero-associative and modular auto-associative memories. We have also tested backpropagation NNs with one hidden layer. Associative memories displayed recognition quality not worse than backpropagation networks.

1. INTRODUCTION

Chemical image (CI) recognition is identification of substance using chemical sensors' data. To solve these problems one generally use classical methods of statistical analysis such as principal component (PCA), discriminant factor analysis et c. Nowadays Feed-Forward neural networks and Kohonen's self-organizing maps are also introduced (applied)\cite{1,2}. In this case one can reduce system tuning to training neural networks using experimental data. Unfortunately, training of neural networks takes a lot of time, so their application is often discouraged. Therefore, associative memories are welcomed because of their greatly faster training (only within one iteration for Hopfield-type networks) \cite{3}. The major disadvantage of associative memories is a rather low generalizing ability that increases sensor stability requirements. But new learning algorithms and architectures of associative memory enable to overcome these difficulties. In particular, modular and hetero-associative memories could be used.

In the case of piezoelectric transducers adsorption process can be directly characterized with the resonance frequency change values (at different moments) and their time derivatives, as well as various combinations of the above values. One should note that use of approaches based on experimental curves fitting with analytical functions to obtain the response parameters seems to be inappropriate in this case. The reason is that very often the simple kinetic models for adsorption cannot adequately describe the curves for the sensor array response, because of Multi-Component Mixture's (MCM) components diffusion into the sensitive coating bulk, change of its structure, uncontrolled fluctuation of temperature and pressure. Indeed, if one uses for MCM CI formation only stationary values (i.e., those that characterize a system after an equilibrium has established), then one may omit important information concerning character of interaction between highly volatile components with different mobilities/activities and the sensitive layer. And the sensitive layer prehistory strongly affects the features of the receptor/analyte interaction in the initial portion of the kinetic curve. Thus the problem how to choose an optimal range of the initial data that could take into account the kinetic and stationary features of the sensor array response, as well as some leveling effects of the pre-starting procedure, are important for both fundamental science and practical applications.

The goal of this paper is to prove possibility of practical application of associative memories to chemical image recognition. We have made a series of experiments on learning to recognize some different substances (perfumes, ethanol, et c.) using associative and Feed-Forward neural networks. The research was done using neural software package NeuroLand designed at Institute of Mathematical Machines and Systems of NASU and the multi-sensor piezo-crystal gas analyzer designed at Institute of Semiconductor Physics of NASU \cite{3}.

2. EXPERIMENTAL TECHNIQUE AND GIVEN DATA

To obtain chemical images we have used an experimental assembly based on Quartz Crystal Microbalance (QCM). The universal neural computer NeuroLand was used to process digitised data: It enables to model different types of neural networks, train and test them.
2.1. The QCM experimental assembly

In sensors based on piezoelectric physical transducers there is the dependence between resonance frequency \( f \) of acoustical resonator and mass \( \Delta m \) of a substance at sensor's surface. It is used to detect intermolecular interactions. For measurements in gas phase dependence between the resonance frequency change \( \Delta f \) and the mass change \( \Delta m \) on its surface is assumed to be linear in accordance with well-known Sauerbreu's equation.

In this work QCM-based arrays were used as prototypes an artificial nose for testing of some perfumes and vapors. For experiments a full-automatic 10-MHz quartz crystals (AT-cut, RK169) based on 8-channel systems with a measurement step of 1 s were used. The e-Nose instrument involves the following units: a thermostatic measuring chamber with a flow-type sensor array; a quartz oscillator unit; a frequency counter with RS232 interface based on AT89C2051 microprocessor; a gas mixtures generator; a computerized system to collect and process information. The general view of the e-Nose instrument is shown in Fig. 1. The sensor units have a specially designed Teflon holder in which the quartz crystals were fixed along its perimeter. So, only one side of quartz crystal was covered by sensing material and contacted with changing ambient. The flow of buffer gas (argon) was used for recovery thin film coating after adsorption phase. It was shown that construction with fixation of quartz crystals along perimeter is preferable due to strong decrease of noise.

The measuring procedure involved the following stages at the temperature 37±0.3 °C: argon circulation until the transducer frequency is stabilized (±2 Hz); vapor-gas mixture circulation at a gas-carrier rate of 20-50 ml/min.; circulation with argon air until the QCM frequency returns to its initial value.

The following five types of perfumes (GOST 17237-93) as typical MCM were tested: eau-de-Cologne “Siren”, “Landish” (produced by “Effect”, Kharkov); eau-de-Cologne “Russkij Les” (OAO, produced by “Kombinat Krimskaja Rosa”, Simferopol), eau-de-Cologne “Shipr” and “Roksolana” (produced by PKK “ROSO”, Zhovka) as well as water and ethyl alcohol. The sample volume was 12 ml in every case.

![Fig. 1. Overall view of QCM experimental appliance](image)

**Fig. 1.** Overall view of QCM experimental appliance

**2.2 Input data for training NN**

For numerical experiments we have used measurements for seven substances (see Table 1).

Most of chemical images contained data for 7 sensors. Time series for each sensor were 286 ticks long. For instance, in the Fig. 2 we can see 7 sensors' output for the odor of ethanol.

Approximately, these regularities are close to exponential. On the other hand, a noise component is rather strong, and starting pieces of time series are sometimes instable. Note that different chemical images for each class are also highly scattered. For certain images intra-class dispersion may exceed 10%.

Data might enter into NN’s input immediately, or after PCA preprocessing. Principal components were computed using the entire training sample. Otherwise NNs were trained and tested using 6 classes only (57 images, classes #1-6).

In this paper we evaluate maximum classification quality using different NNs. Also we try to find optimal

![Fig. 2. Response of 7 sensors (A, B, D-H) for ethanol's smell](image)

**Fig. 2.** Response of 7 sensors (A, B, D-H) for ethanol's smell

**Table 1. Classes of chemical images**

<table>
<thead>
<tr>
<th>Class #</th>
<th>Substance</th>
<th>Number of Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ethanol</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>“Landish”</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>“Russki Les”</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>“Roksolana”</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>“Shipr”</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>“Siren”</td>
<td>15</td>
</tr>
<tr>
<td>7</td>
<td>Water</td>
<td>3</td>
</tr>
</tbody>
</table>

2631
parameters of preprocessing technique. Sample window metrics (size and starting point) and number of PCs were such tuneable parameters.

The "correctly recognized images"/"all test images" ratio is called classification rate. Each value of classification rate was obtained in averaging over several "training/test" dissections of the sample. These dissections were randomly generated holding all other experiment conditions constant. To determine recognition quality we have used 50 for the modular associative classifier, and 10 dissections for the hetero-associative one. For the Feed-Forward NN recognition quality was estimated in averaging over 10 results with different initial states of the network.

3. ARCHITECTURE AND ALGORITHMS OF NEURAL NETWORKS

3.1. Associative memories

Associative memories are preferred due to fast learning done within single iteration in computing all neurons' weights. Unfortunately, most common models like Hopfield networks or bi-directional associative memories deal with bipolar data only [4]. It's needed to use convergence examination for restoration of corrupted data. But we need only classify chemical sensors' data. So, we should build an associative-memory classifier, which will use real data vectors. We have considered two associative memory classifiers: hetero-associative and a modular one.

3.1.1. Hetero-associative classifier

Hetero-associative classifier is a single-layered network with linear activation function; chemical sensors' data are directed to its N inputs. The quantity of neurons K is equal to number of classes (chemical substances). Weights are computed using pseudo-inverse learning rule [6]. These rules cold be obtained solving stability equation:

$$Bx_i = y_i$$

where: B is a KxN weight matrix of thee network, x_i - N-dimensional input vector, y_i - K-dimensional output vector, its k-th component is equal to +1 if corresponding input belongs to class k, the rest of components are -1.

The training data array could be represented as two matrices: X contains M data vectors, and Y contains M desired network's responses.

Solution of (1) is:

$$B = YX^+$$

where X^+ is a pseudo-inverse matrix to X. X^+ is usually computed using Greville's formulae [6].

3.1.2. Modular auto-associative classifier

It consists of K neural modules; each of them is a conventional auto-associative network with real stored images. Every module memorise N-dimensional data vectors for certain class. Data are memorised in k-th module computing projection matrix.

$$C_k = X_kX_k^*$$

where X_k is a matrix containing data vectors for k-th class.

During recognition of an input vector z the post-synaptic potential is computed:

$$s_k = C_kz$$

The letter is a projection of an input vector to linear span of training images for k-th class. The decision on belonging of input image is made maximizing scalar product:

$$q_k = s_kz^T$$

where z^T is a conjugate vector to z.

The vector z is treated as belonging to k* that have q_k* → max.

3.2. Classifier using Feed-forward NN

Nowadays feed-forward neural networks and backpropagation learning algorithms are conventional techniques for chemical image recognition [1,2]. Strength and weaknesses of these NNs are well known; therefore we used such networks, basically, as a benchmark for associative-memory classifiers.

We have used NNs with one hidden layer. Hidden neurons had a sigmoid activation function, without a bias. Number of inputs is a dimension of input vector; for each class there is one output neuron. Data were normalized to [0, 1] range. We have tested networks with 5, 10 and 15 hidden neurons. Learning algorithm was a standard online backpropagation with learning rate 0.01 and momentum 0.15. Learning took 6000 epochs (about 10 min. of CPU time for Intel Celeron 600 MHz); so mean square error could be decreased to 10^{-7} - 5⋅10^{-5}.

4. CLASSIFICATION QUALITY DEPENDING ON SIZE OF TRAINING SAMPLE

For these series of experiments we have used raw data for all available ticks from 1st to 286th. In the case of usage of principal components the 50 highest ones were entered into classifier. Learning took about 4 sec. of CPU time for Intel Celeron 600 MHz. Results for modular and hetero-associative classifier are shown in Fig. 3.
module memorizes images of one class only; so there was no overflow

Fig. 3. Classification quality subject to training image number

Fig. 4. Recognition quality subject to initial measurement and number of ticks used for modular network

Fig. 5. Recognition quality subject to initial measurement and number of ticks used for hetero-associative network

5. DEPENDENCES ON WINDOW SIZE

These experiments were dedicated to find optimal parameters of data window for entering data into network. All types of NNs were tested. Modular network was trained using 4 examples for each class; training sample for hetero-associative one took 60% of available images. Positions of initial and final time ticks were scanned with step of 15 ticks.

Dependences of classification quality on size of data vector for associative memories are shown in the Fig. 4, 5. Such dependences for Feed-forward NNs are displayed in Fig. 6. Starting point of a windows serves as a parameter in all the figures.

Note that both classifiers achieve best quality in excluding starting segment of data (100-150 ticks) and taking not more than 150-200 points for the network input. This may be explained taking into account instability of starting part of curves shown in fig 2. Note also that hetero-associative memory displays slightly better quality than modular one (73% vs. 71%). This might be caused only by experimental error.

Results for Feed-Forward networks show that initial parts of sensor response curves have low information. Best classification rates - 75-77% - were obtained excluding first 100 ticks and taking data for 50-100 ticks of the middle of response curves.

6. CLASSIFICATION RATE SUBJECT TO NUMBER OF PRINCIPAL COMPONENTS

These experiments were dedicated to find minimal quantity of PCs required to keep admissible classification rat for associative memories. Modular network was trained using 4 examples for each class; training sample for hetero-associative one took 60% of available images. Test results are shown in Fig. 7.
7. CONCLUSION

Experimental results described above show that associative memories provide approximately the same recognition quality as feed-forward neural networks. We stress that non-iterative nature of neural associative memories makes them quite attractive. This is not only high-speed learning but also learning independent on a (random) network initialization. Even a low generalisation ability of associative memories may be turned into advantage; using more stable chemical sensors they will be able to recognise thousands sorts of chemical substance.

8. ACKNOWLEDGEMENTS

This work was partially supported by INTAS grant No. 01POLL-0257.

Calyxarenes were kindly given to us by Prof. V.I. Kal'chenko (Institute of Organic Chemistry of NASU, Kyiv). Linear polyacenes were kindly given to us by Prof. Ya.I. Vertsimakha (Institute of Physics of NASU, Kyiv). Myro cyclic compounds were kindly given to us by Prof. Ya.D. Lampeka (Institute of Physical Chemistry of NASU, Kyiv).

9. REFERENCES