An eNose algorithm for identifying chemicals and determining their concentration

L. Carmel[a,*], N. Sever[a], D. Lancet[b], D. Harel[a]

[a]Department of Computer Science and Applied Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel
[b]Department of Molecular Genetics, The Weizmann Institute of Science, Rehovot 76100, Israel

Abstract

We propose an algorithm for use with multisensor systems that is capable of the following: (a) identify an analyte independently of its concentration; (b) estimate the concentration of the analyte, even if the system was not previously exposed to this concentration; (c) tell when an analyte is of a chemical type not previously presented to the system. The algorithm, based upon recent work of Hopfield, uses the multiplicity of sensors explicitly, and is intuitive and easy to implement. We have tested it against real data, and it exhibits high quality performance.
© 2003 Elsevier Science B.V. All rights reserved.

Keywords: Electronic noses; Classification; Identification; Concentration estimation; Reject option; Hopfield algorithm

1. Introduction

An electronic nose (eNose) is an analytic device used for detecting vapor chemicals [1]. It functions using an array of broadly tuned chemical sensors, i.e. sensors that interact with a broad range of chemicals with varying strengths. Consequently, an incoming analyte stimulates many of the sensors in the array, and elicits a characteristic response pattern. These patterns are then further analyzed for the benefit of the specific application.

Classification tasks, designed to determine the identity of incoming analytes, are by far the most popular form of analysis. Classification is usually achieved by one of the well established classical algorithms known from the field of pattern recognition, such as the k-nearest-neighbors, the shortest Mahalanobis distance, the linear and quadratic discriminant analysis, and the multilayer perceptron [2]. There are many applications of such algorithms, e.g. for quality assessment of food products [3,4], for medical diagnostics [5], and even in the automobile industry [6].

A classification algorithm designed for a specific problem has to go through a preliminary *training phase*, in which it studies the patterns typical to the problem. During this phase, the algorithm is fed with samples whose classification is known in advance. When this phase is finished, the algorithm can be used to identify samples whose classification is not known. Obviously, the algorithm can only identify the classes that were presented to it during the training phase. In some cases, it can also *reject* samples, thus determining that they do not belong to any of the classes it has 'seen'.

This methodology is good as far as it goes. Its main weakness is in lacking prediction ability. It cannot infer anything positive regarding information that was not present during the training phase. We demonstrate this by an example. Suppose that during the training phase we used samples of chemicals a and b in concentrations 100 and 70 ppm, respectively. Introducing chemical a in concentration 100 ppm to the classification algorithm should not lead to any special difficulties. However, introducing chemical a in concentration 70 ppm, or chemical b in concentration 120 ppm, would often reduce the performance of the algorithm. Changing the concentration influences the patterns, and classification algorithms exhibit increased misclassification rates.

The susceptibility of the patterns to concentration poses unique challenges in the usage of eNoses. An ideal classification algorithm should cope with this by having the following two properties:

1. It should identify chemicals independently of their concentration. This is especially important in applications

* Corresponding author.
E-mail addresses: liran.carmel@weizmann.ac.il (L. Carmel),
doron.lancet@weizmann.ac.il (D. Lancet),
dharel@weizmann.ac.il (D. Harel).

where we do not have full control over the environmental conditions during the measurement.

2. It should be able to estimate the concentration of the analyte. This is required in applications where the concentration has some meaning. For example, if we use an eNose to alert us to a certain chemical becoming too concentrated in the surroundings.

The need to deal with the concentration dependency of eNoses has been long recognized. One possible approach to dealing with this problem is to treat different concentrations of the same chemical as if they were different chemicals. For example, if chemical *a* is expected to be found within the concentration range 70–130 ppm, we may train the algorithm on, say, the three concentrations 80, 100 and 120 ppm, and then classify new samples to one of these three classes. A nice example of such a study can be found in [7]. The advantage of this approach is that it leaves us in the well-studied framework of regular classification theory. However, its use is limited to cases where the number of chemicals to be classified is small, and when a crude estimation of their concentrations suffices. If finer predictions are required, this approach would turn the number of classes to be trained upon impractically large. A better approach would be to design an algorithm that has concentration prediction capabilities. A few attempts in this direction have been reported, usually using various kinds of relatively complex artificial neural networks; see, e.g. [8–10].

In this paper, we propose an algorithm that is able to identify chemicals independently of their concentration, as well as to quantify their concentration, even if the particular concentration in question was not present during the training phase. The algorithm is distinct for its intuitive approach, and utilizes in a very straightforward way the special properties of a multisensor system. Besides its concentration prediction capabilities, the algorithm is equipped with a reject option, and is thus able to tell when a sample is not one of the chemicals upon which it was trained. The algorithm is inspired by the work of Hopfield [11], see next section, who proposed a similar one to explain how olfactory data analysis is carried out in the brain.

2. The algorithm

Hopfield [11] proposed an algorithm in an attempt to explain how the brain deals with olfactory processing tasks such as identifying chemicals, separating signal from background, and determining the components of a mixture. The mammalian olfactory system, including that of human, consists of 300–1000 different types of olfactory receptors [12,13]. The Hopfield algorithm explicitly uses this multiplicity of receptors, by separately querying each of the receptors, and making decisions based on all the answers received. Hopfield assumed the simplest kind of chemical interaction between an olfactory receptor and an odor

ligand, and so he was able to obtain a simple description of the concentration-dependency of each individual olfactory receptor. This response function was invertible, and Hopfield formulated a closed-form expression for the concentration as a function of the response. Finding the odor-specific parameters of this closed-form expression is equivalent, á la Hopfield, to the learning process carried out in the brain upon sniffing.

We apply similar ideas to eNoses, where the sensors serve as the analogs of the biological receptors. However, unlike Hopfield, we do not try to find an analytic expression for the concentration-dependency of the sensors, but rather measure these experimentally and invert them by interpolation. Hopfield had demonstrated his ideas by simulating the olfactory system, which comprises hundreds of receptors. We show that in eNose systems, despite the fact that the number of sensors (typically 10–30) is significantly smaller, the same principles can be used to construct a powerful and reliable algorithm.

Assume our eNose consists of m different sensors, and suppose that it can be exposed to any of n different analytes, where the concentration of the ith analyte c_i can be anywhere in the range $c_i^{\min} \leq c_i \leq c_i^{\max}$ (where $c_i^{\min}$ and $c_i^{\max}$ are odor-specific predetermined constants). We shall hereinafter call these n analytes *candidates*. Upon the introduction of a new sample, our algorithm is able to do the following:

(1) If the sample is one of the n candidates, the algorithm identifies it and estimates its concentration.
(2) If the sample is not one of the n candidates, the algorithm rejects it (i.e. it is able to determine that the sample is from a substance it is unfamiliar with).

2.1. The training phase

Each of the n candidates should be presented to the eNose in several different concentrations, and the responses are recorded. The data is then used to construct the $n \times m$ functions

$$r_{ij} = f_{ij}(c_j), \tag{1}$$

where r_{ij} is the response of the ith sensor to the jth candidate in concentration c_j, for $i = 1, \ldots, m$, $j = 1, \ldots, n$, and $c_j^{\min} \leq c_j \leq c_j^{\max}$. Our algorithm requires $f_{ij}(c_j)$ to be monotonic in the range $c_j^{\min} \leq c_j \leq c_j^{\max}$, so that it has a well-defined inverse. Otherwise, there are no constraints on its shape.

These functions can be evaluated using diverse techniques. If the response is believed to be linear with respect to the concentration, a linear regression can be used. If some other analytic shape of the response is speculated, a standard curve-fitting process can be carried out. Otherwise, interpolation algorithms can be applied, such as linear interpolation, or piecewise cubic spline interpolation. In Section 4, we show some examples of functions $f_{ij}(c_j)$ that were obtained using piecewise cubic spline interpolation.

2.2. Classification at work

Being invertible, the functions $f_{ij}(c_j)$ turn each of the sensors into a kind of concentration sensor. That is, if a sample of candidate j is measured and elicits a response r_i in sensor i, then this sensor estimates the concentration of the sample by

$$c_{ij} = f_{ij}^{-1}(r_i). \tag{2}$$

If we have m different sensors, we get m different estimations of the concentration, $c_{1j}, c_{2j}, \ldots, c_{mj}$, and we may supply an overall estimate by taking the average,

$$\bar{c}_j = \frac{1}{m}(c_{1j} + c_{2j} + \cdots + c_{mj}). \tag{3}$$

The crucial point here is that if the sample is indeed the jth candidate, the series of estimates $c_{1j}, c_{2j}, \ldots, c_{mj}$ will be all close to each other, but if the sample is not the jth candidate, the sensors would produce an array of totally different estimates. We can, therefore, use the variance

$$\sigma_j = \frac{1}{m-1}[(c_{1j} - \bar{c}_j)^2 + (c_{2j} - \bar{c}_j)^2 + \cdots + (c_{mj} - \bar{c}_j)^2], \tag{4}$$

to decide the identity of the sample—if σ_j is small enough (say lower than some threshold σ_0), then the sample is indeed the jth candidate. But, if it is large, the sample must be different.

The resulting algorithm is described schematically in Fig. 1, in its simplest form. We have made the algorithm more robust by adding to it two more components:

(1) Let $r_{ij}^{\max} = f_{ij}(c_j^{\max})$ be the maximal possible response of sensor i when exposed to candidate j whose concentration is known to be less than $c_j^{\max}$. If a sample elicits a response $R_i > r_{ij}^{\max}$, we set $c_{ij} = $ INVALID, meaning that the estimate c_{ij} cannot be reliably obtained.
(2) The functions $f_{ij}(c_j)$ are all very close in values for low concentrations, since they all should be approximately zero when the concentration tends to zero. Hence, when R_i is small enough, σ_j will be small for any j, and the reliability of the algorithm would be inadequate. This is just an expression of the well-known fact that each sensor has a detection threshold, below which it cannot be reliably used. It is, therefore, important to set $c_j^{\min}$ above the detection threshold of the sensors. In analogy with the previous item, we define $r_{ij}^{\min} = f_{ij}(c_j^{\min})$ to be the minimal possible response of sensor i when exposed to candidate j whose concentration is known to be higher than $c_j^{\min}$. If a sample elicits a response $R_i < r_{ij}^{\min}$, we set $c_{ij} = $ INVALID, meaning that the estimate c_{ij} cannot be reliably obtained.

When calculating the numbers c_{ij}, we also count how many INVALIDs are obtained for each candidate. If for candidate j

```
for   j = 1, ..., n
   assume that the sample is the j'th candidate;
   calculate c̄_j and σ_j
end for
choose r as the index of the "best" candidate, r = arg min_j σ_j
if σ_r < σ_0
   identify the sample as being of the type of the rth candidate;
   estimate its concentration as c̄_j
else
   reject the sample
end if
```

Fig. 1. Schematic flow of the algorithm. Here, n is the number of candidates, and σ_0 is a predetermined threshold value.

we have counted too many of these (typically, more than 2), we decide that the sample cannot be this candidate.

3. Experimental

The algorithm was tested against data we have collected using a MOSESII eNose [14] with two sensor modules: an eight-sensor quartz-microbalance (QMB) module, and an eight-sensor metal-oxide (MOX) module. Our eNose, therefore, consists of $m = 16$ sensors. The samples were put in 20 ml vials in an HP7694 headspace sampler, which heated them to 40 °C and injected the headspace content into the electronic nose in a flow of 25 ml/min. There, the sample was first introduced into the QMB chamber, whence it followed to the 300 °C heated MOX chamber. The injection lasts for 20 s (transient signals), and is followed by a 15 min purging stage using synthetic air.

In a single measurement, a sensor produces a signal over time. Hereinafter, we define the response of a sensor in the traditional way of taking the difference between the maximum of the signal and its baseline, although there are more sophisticated approaches to capturing the information in such signals; see, e.g. [15].

We collected two different datasets:

- The *training dataset* was constructed from $n = 10$ candidates, each measured in six different concentrations.[1] The candidates and their concentrations were chosen so that they all have comparable ranges of response. The choice of 10 candidates seems to be rather representative, since typical eNose applications, reported upon in the literature,

---

[1] Except for 2,3-butanedione, which was measured only in four concentrations, and butyl butyrate, which was measured in five concentrations.

Table 1
The candidates and their concentrations; molar fractions are measured in PEG-400 solution

Candidate	Concentrations measured (molar fraction)					
1-Methylpyrrole	0.0908	0.1665	0.2306	0.2855	0.3331	0.3747
1-Propanol	0.1055	0.1909	0.2614	0.3206	0.3710	0.4144
2,3-Butanedione	0.0918	0.1316	0.1681	0.2016	–	–
2,6-Dimethylpyridine	0.0711	0.1328	0.1867	0.2344	0.2768	0.3147
2-Methyl-2-pentenal	0.0721	0.1345	0.1890	0.2371	0.2797	0.3179
4-Methylanisole	0.0657	0.1233	0.1742	0.2195	0.2601	0.2967
Amyl formate	0.0632	0.0919	0.1189	0.1443	0.1683	0.1910
Butyl butyrate	–	0.0983	0.1406	0.1791	0.2142	0.2465
Isoamyl formate	0.0633	0.0920	0.1190	0.1445	0.1685	0.1912
Toluene	0.0770	0.1112	0.1430	0.1726	0.2002	0.2260

are usually required to distinguish between only a small number of stimuli types. Each sample was diluted in poly ethylen glycol 400 (PEG-400), and the concentrations were measured in molar fractions (i.e. the ratio between the number of moles of the sample and the total number of moles in the solution). The candidates and their measured concentrations are listed in Table 1. A candidate in a certain concentration was measured in batches of at least four successive measurements. In total, this dataset consists of 269 measurements.

- The *validation dataset* was used to test the performance of the algorithm. It consisted of three different types of measurements:
  (a) Six candidates were chosen at random and measured at concentrations (chosen again at random) from within those that were used during the training phase.
  (b) Seven candidates were chosen at random and measured at concentrations not present during the training phase.
  (c) Four non-candidate chemicals were measured at seven different concentrations.

The list of measured chemicals and their concentrations is shown in Table 2. Every chemical in every one of the concentration was measured twice. In total, this dataset consists of 40 measurements.

4. Results

4.1. The training phase

In this phase, we construct the functions $f_{ij}(c_j)$ using the data collected in the training dataset. Each candidate j was measured in six different concentrations that we denote by $c_j^1, \ldots, c_j^6$. The corresponding responses of the ith sensor, denoted $r_{ij}^1, \ldots, r_{ij}^6$, are calculated as the average of the repetitions. The six values $(c_j^1, r_{ij}^1), \ldots, (c_j^6, r_{ij}^6)$ are the data from which we have to evaluate $f_{ij}(c_j)$. This we did by using the piecewise cubic spline interpolation of Matlab®, although other kinds of interpolation can be used. Two examples of what these functions look like are shown in Fig. 2. In some cases, the concentration dependency of the response is very close to linear, as illustrated in Fig. 3. In such cases, $f_{ij}(c_j)$ can be approximated well by linear regression on the data.

4.2. Validation against the training dataset

An initial examination of the algorithm would be to test it against the same dataset that was used to construct it—the training dataset. Here, we took the threshold value to be $\sigma_0 = 0.03$ molar fraction, and obtained the following results:

- All the samples were classified correctly (i.e. 100% correct classification).

Table 2
The chemicals measured in the validation set and their concentrations; molar fractions are measured in PEG-400 solution

| Group	Candidate	Concentration (molar fraction)
(a)	1-Methylpyrrole	0.1665
(a)	2,3-Butanedione	0.2016
(a)	2,6-Dimethylpyridine	0.0711
(a)	4-Methylanisole	0.2195
(a)	Amyl formate	0.0919
(a)	Butyl butyrate	0.1406
(b)	1-Methylpyrrole	0.1998
(b)	1-Propanol	0.1503
(b)	2,6-Dimethylpyridine	0.2113
(b)	2-Methyl-2-pentenal	0.2590
(b)	4-Methylanisole	0.0954
(b)	Amyl formate	0.0778
(b)	Isoamyl formate	0.1319
(c)	Hexyl alcohol	0.1244
(c)	Isoamyl propionate	0.0967
(c)	Isoamyl propionate	0.1764
(c)	Heptyl alcohol	0.1116
(c)	Heptyl alcohol	0.2389
(c)	*cis*-3-Hexenyl acetate	0.1009
(c)	*cis*-3-Hexenyl acetate	0.2519
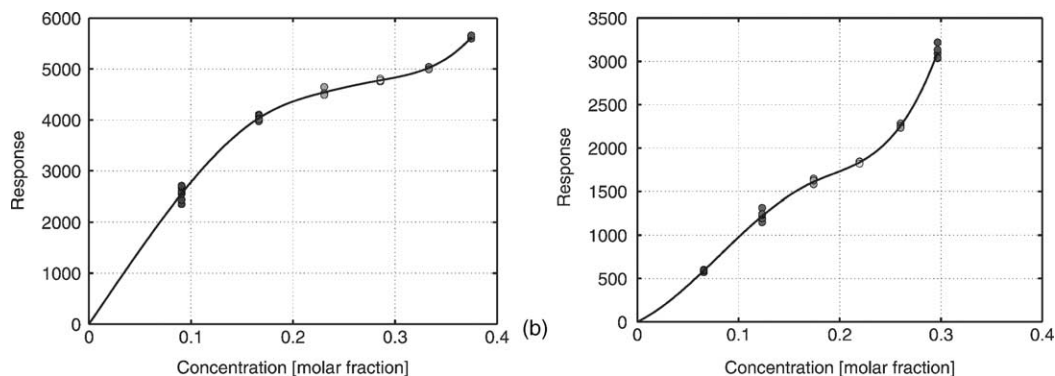
Fig. 2. Two examples of the functions $f_{ij}(c_j)$. The dots are the measurements, and the solid line is the model built by piecewise cubic spline interpolation. The left-hand figure depicts the response of one MOX sensor as a function of 1-methylpyrrole concentration, and the right-hand figure depicts the response of one QMB sensor as a function of 4-methylanisole concentration. Linearity is not important, as long as the curve is monotonic. Molar fractions are measured in PEG-400 solution.
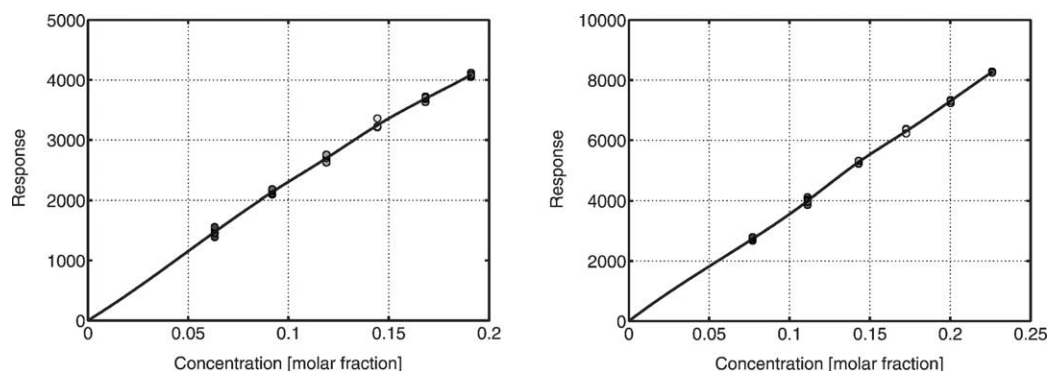

Fig. 3. Two examples where the functions $f_{ij}(c_j)$ are close to linear. The dots are the measurements, and the solid line is the model built by piecewise cubic spline interpolation. The left-hand figure depicts the response of one MOX sensor as a function of amyl formate concentration, and the right-hand figure depicts the response of one QMB sensor as a function of toluene concentration. Molar fractions are measured in PEG-400 solution.

- The concentrations were predicted with extremely high precision, the average relative error being 1.4%.

4.3. Validation against the validation dataset

Applying the algorithm to the validation dataset, which is, of course, a far more difficult task, we still saw a very high level of performance. In total, 95% of the samples were correctly classified, with an average relative error of 10% in prediction the concentration (again, we used $\sigma_0 = 0.03$).

Here, is a more detailed discussion of the performance with respect to the three aforementioned groups of measurements:

(a) The samples were candidates at concentrations identical to those presented during the training phase. All samples were classified correctly, giving 100% accurate classification. Concentrations were predicted with a precision of 7.8% average relative error. Naturally, the concentration predictions are slightly inferior with respect to those obtained for the training dataset. This is perfectly reasonable, considering that eNose measurements are reproducible up to small drifts in time. When

the same sample is measured at two different times, small changes in the patterns are usually observed. Nevertheless, our results show that the repeatability is good enough to allow for a robust algorithm.

(b) The samples were candidates at concentrations not present during the training phase. Again, 100% of the samples were correctly classified (which is inarguably the most exciting result we have obtained). Concentrations were predicted with a precision of 11.8% average relative error.

(c) The samples were chemicals not present at all during the training phase. Taking a rejection as the correct classification, two measurements were misclassified, giving an 83.3% rate of correct classification. The two samples of isoamyl propionate at a concentration of 0.18 molar fraction were erroneously classified as butyl butyrate.

5. Summary and discussion

We have presented an algorithm, inspired by the work of Hopfield on olfaction, that identifies chemicals presented to an eNose, regardless of their concentration. Moreover, the

algorithm can estimate with good precision the concentration of the sample, and it can detect situations where the sample is a chemical that has never been presented to the system before. The algorithm is very intuitive and easy to implement, and exhibited good performance when applied to real data.

The algorithm is not limited to eNose systems; it can, in fact, be applied to any multisensor system in which the sensors respond monotonically to the stimuli. Our experiments demonstrate that 16 sensors are enough for obtaining good results. As the number of sensors in the system grows, performance is expected to be improved. The robustness of the algorithm is also expected to improve with the increase in the number of sensors. Actually, one can think of the response of each sensor as a random variable, from which we can estimate the concentration. As is the situation in statistical inferences, the estimator (concentration) will be more accurate as the number of samples (sensors) increases.

One of the attractive features of the algorithm is its simplicity. Algorithmically speaking, it is clean and simple, having essentially a single free parameter, σ_0. We anticipate, however, that the performance can be further increased if additional parameters are introduced. We may use a candidate-specific σ_0, that is, candidate j will be rejected if $\sigma_j > \sigma_0^j$, $j = 1, \ldots, n$, with σ_0^j predetermined constants. In this way, differences in the "natural" variance of different chemicals may be incorporated. A further generalization of this idea calls for using a candidate-specific concentration-dependent σ_0. Thus, a candidate j will be rejected if $\sigma_j > \sigma_0^j(\bar{c}_j)$, $j = 1, \ldots, n$, with $\sigma_0^j(\bar{c}_j)$ predetermined functions. These functions might be constructed during the training phase similarly to the f_{ij}'s.

We would like to point out that our algorithm seems to be robust with respect to noises. Since inference is done from each sensor individually, white noises are expected to be "averaged out".

6. Conclusions

The suggested algorithm was demonstrated to have excellent performance when applied to our datasets. This should serve to convince other researches to see whether it could be useful for related problems.

The spectrum of applicability of this algorithm appears to be very broad. It can be used whenever there is a need to distinguish between different types of stimuli in an uncontrolled environment, or whenever there is a need to determine concentrations, as long as the sensing mechanism is multidimensional.

References

[1] J.W. Gardner, P.N. Bartlett, Electronic Noses, Principles and Applications, Oxford University Press, New York, USA, 1999.

[2] A. Webb, Statistical Pattern Recognition, Arnold, London, Great Britain, 1999.

[3] S. Hahn, M. Frank, U. Weimar, Rancidity investigation on olive oil: a comparison of multiple headspace analysis using an electronic nose and GC/MS, in: Proceedings of the 7th International Symposium on Olfaction and Electronic Nose, ISOEN 2000, pp. 49–50, July 2000.

[4] C. Di Natale, A. Macagnano, E. Martinelli, R. Paolesse, E. Proietti, A. D'Amico, The evaluation of quality of post-harvest oranges and apples by means of an electronic nose, Sens. Actuators B 78 (2001) 26–31.

[5] Y.-J. Lin, H.-R. Guo, Y.-H. Chang, M.-T. Kao, H.-H. Wang, R.-I. Hong, Application of the electronic nose for uremia diagnosis, Sens. Actuators B 76 (2001) 177–180.

[6] A. Guadarrama, M.L. Rodriguez-Mendez, J.A. de Saja, Conducting polymer-based array for the discrimination of odours from trim plastic materials used in automobiles, Anal. Chim. Acta 455 (2002) 41–47.

[7] S.J. Qin, Z.J. Wu, A new approach to analyzing gas mixtures, Sens. Actuators B 80 (2001) 85–88.

[8] B.-S. Joo, N.-J. Choi, Y.-S. Lee, J.-W. Lim, B.-H. Kang, D.-D. Lee, Pattern recognition of gas sensor array using characteristics of impedance, Sens. Actuators B 77 (2001) 209–214.

[9] R.M. Negri, S. Reich, Identification of pollutant gases and its concentrations with a multisensor array, Sens. Actuators B 75 (2001) 172–178.

[10] J. White, T.A. Dickinson, D.R. Walt, J.S. Kauer, An olfactory neuronal network for vapor recognition in an artificial nose, Biol. Cybern. 78 (1998) 245–251.

[11] J.J. Hopfield, Odor space and olfactory processing: collective algorithms and neural implementation, Proc. Natl. Acad. Sci. USA 96 (1999) 12506–12511.

[12] L. Carmel, D. Harel, D. Lancet, Estimating the size of the olfactory repertoire, Bull. Math. Biol. 63 (2001) 1063–1078.

[13] D. Lancet, N. Ben-Arie, Olfactory receptors, Curr. Biol. 3 (1993) 668–674.

[14] J. Mitrovics, H. Ulmer, U. Weimar, W. Gopel, Modular sensor systems for gas sensing and odor monitoring: the MOSES concept, Acc. Chem. Res. 31 (1998) 307–315.

[15] L. Carmel, S. Levy, D. Lancet, D. Harel, A new feature extraction technique for electronic noses, in: Proceedings of the 9th International Meeting on Chemical Sensors (ICS2002), Boston, USA, 7–10 July 2002 93 (2003) 67–76.

Biographies

L. Carmel received his BSc in physics at Tel-Aviv University, Israel, in 1991, and his MSc degree in physics at the Technion—Israel Institute of Technology, in 1998. He is currently completing his PhD studies in the Department of Computer Science and Applied Mathematics at the Weizmann Institute of Science, Israel. His research deals with materializing odor digitization, transmission and reproduction, and it involves many kind of mathematics (e.g. multivariate data analysis, statistical pattern recognition), biology (e.g. the sense of smell, receptor repertoires), and chemistry (e.g. electronic noses, chemical sensors).

N. Sever received her BSc degree in Plant Sciences at the Hebrew University of Jerusalem, Israel, in 2000. She is currently completing her MSc studies in the Department of Field Crops, Vegetables and Genetics at the Hebrew University. Her research deals with molecular biology and genetic engineering in medicago truncatula. She also works in the electronic nose laboratory of the Weizmann Institute of Science.

D. Lancet is the Ralph and Lois Silver Professor of Human Genomics at the Department of Molecular Genetics of the Weizmann Institute of Science, Rehovot, Israel. He has been Head of the Crown Human

Genome Center at Weizmann since 1998. Prof. Lancet received his BSc degree in chemistry at the Hebrew University of Jerusalem in 1970 and his PhD degree in chemical immunology at the Weizmann Institute in 1978. He headed Weizmann's Department of Membrane Research and Biophysics (1995–1997). Lancet pioneered genome research in Israel, and currently operates Israel's National Laboratory for Genome Infrastructure. His research interests include the genetic basis of the sense of smell and of inherited diseases, formalisms of molecular recognition and computer models for the origins of life. As part of an extensive involvement in the bioinformatics scene, his team has developed GeneCards, a world-known compendium of human genes, and GeneNote-a whole-genome DNA array register. He received, among others, the First Takasago Award of the American Association for Chemoreception Sciences (1986), and the R.H. Wright Award in Olfactory Research (1998). Lancet is a member of the European Molecular Biology Organization since 1996.

D. Harel is the William Sussman professor of mathematics at The Weizmann Institute of Science in Israel, and has been Dean of the Faculty of Mathematics and Computer Science there since 1998. He is also co-founder of I-Logix, Inc., Andover, MA. He received his BSc from Bar-Ilan University in 1974, his MSc from Tel-Aviv University in 1976, and his PhD from MIT in 1978. He has worked in several areas of theoretical computer science, including computability, finite model theory and logics of programs, and in recent years has become involved in other areas, including software and systems engineering, visual languages, graph layout, modeling and analysis of biological systems, and smell communication. He is the inventor of statecharts, and co-inventor of live sequence charts (LSCs), and was part of the team that designed the Statemate and Rhapsody tools. He has received a number of awards, including ACM's Karlstrom Outstanding Educator Award in 1992. His latest books are "Dynamic Logic" (with Kozen and Tiuryn), MIT Press, 2000, and "Computers Ltd.: What They Really Can't Do", Oxford, 2000.