Multiclass SVMs for Olfactory Classification

CS229 Final Project Report

Michael Styer

Computer Science Department, Stanford University

1 Introduction

The study of olfactory classification has many attractions, philosophical and theoretical as well as practical and functional. As part of the study of human cognition, it is widely recognized that the olfactory sense is a fundamental component of the human perceptual system. The human brain is believed to have evolved around the olfactory sense, and consequently the algorithms evolved for olfaction form the foundations of the algorithms used by the other senses. Input from all sensory modalities converges in the entorhinal cortex, indicating a degree of common processing function across the senses [14]. Olfaction is also an important and valuable perceptual mechanism for making distinctions about objects in and states of the physical world. The development of the so-called "electronic nose" over the past thirty years has led to a wide variety of applications, from quality analysis in the food and pharmaceutical industries to explosive and biological agent screening in security and defense applications. More potential applications present themselves, dependent on advances in sensor technology and signal processing algorithms, among them medical diagnosis, environmental quality monitoring, autonomous bomb disposal robot navigation, and smart fire alarm and industrial monitoring systems, among others.

This research project concentrates on the problem of classification of an unknown odorant into one of a set of previously observed classes, the problem in general of multiclass classification in a supervised learning framework. In the olfactory sensing context there are several components required of a sucessful classification algorithm: *preprocessing*, *feature extraction*, and *classification*. Previous research has applied various statistical learning techniques to each of these sub-problems. For preprocessing, techniques include feature normalization and baseline subtraction; for feature extration, methods include principal component analysis, linear discriminant analysis, and Sammon and Kohonen maps; and for classification, techinques include Bayesian classifiers, K-nearest neighbors, artificial neural networks, radial basis functions and, recently, support vector machines.

The paper proceeds in the standard fashion. In Section 2, we describe the problem of olfactory sensing in general and describe the distinguishing characteristics of the particular problem considered in this research paper. In Section 3, we identify the major trends in previous and current research into this problem. In Section 4, we describe the approach taken in this paper and locate our techniques within the broader scope of related research. In Section 5 we describe the results obtained , and finally in Section 6 we discuss the results and outline directions for further research.

2 Problem Description

The problem of automatic olfactory classification has several related components in addition to the elements of the classification algorithm mentioned above. Perhaps most critical is the sensing apparatus itself. In the mid- to late-1980s, researchers at the University of Warwick devloped sensor arrays with sufficiently broad sensitivity to be used for odor classification as opposed to detection of specific odorants [12, 37, 38]. Gas sensor techonology in general relies on a change in electrical or other properties of some material in the presence of volatile organic compounds (VOCs). Early sensors included inorganic semicoductors (e.g [42]), organic semiconducting polymers [12], and Langmuir-Blodgett films (see, e.g., [2] and [43]). More recent sensor technology includes piezoelectric sensors, which operate by changing resonating frequency due to absorption of VOCs; metal-oxide-silicon field-effect-transistor (MOSFET) devices, which employ a catalytic

metal whose reaction products cause a change in electrical conductivity; and optical fiber sensors, in which a glass fiber is coated with a chemically active material that responds to the presence of VOCs by changing its fluorescent emission spectrum [31]. Cutting-edge research, however, employs nanotubes and other techniques from nanotechnology to achieve substantially better results than previous methods (see, e.g. [41], [3], [25], [46]).

Given an adequate sensor array, olfactory object classification presents several connected challenges [20]. First, nearly all existing sensor types are subject to drift, a slow, random decrease in sensor response when exposed to the same odorants under the same or similar conditions. As a result, sensor patterns learned previously become obsolete, and the system's ability to identify known odors is compromised. The primary solution to this problem has been periodic recalibration with a reference gas, but recently work has been doen to develop signal processing techniques that can remove the dependence on recalibration [8]. Second, due to the omnipresence of VOCs in the environment, and particularly in environments such as food-processing plants or medical facilities in which many of the most promising applications of olfactory sensing technology would be located, any olfactory sensing system must be able to extract the signal of the target olfactant from a broad-spectrum background. Techniques for extracting the signal from the background include simple differencing, relative extraction and fractional extraction.

Third, a sensor response vector should be normalized for processing by subsequent classification algorithms, to eliminate sensitivity to concentration in the classification system. This can be done locally, by normalizing each feature vector independently, or globally, by restricting or scaling the response of each sensor across all feature vectors. Fourth, a multiple sensor array yields a high-dimensional feature space, which may be computationally difficult or impossible to process in the classification stage. As a result, a strategy of feature extraction (e.g. linear discriminant analysis or principal component analysis), or feature subset selection (e.g. randomized search over feature subsets) may be used to choose a subset of the full feature set that gives a good classifier while reducing the problem complexity.

Finally, given a normalized dataset with a selected feature subset, the problem remains to make a classification of an unknown odorant given previous exposure to a set of known odorants. This is a typical object classification problem not dissimilar to text classification or data mining, although the nonlinearity of the sensor responses presents a challenge. Most of the common classification algorithms have been applied to this problem, including linear and quadratic classifiers, K-nearest neighbors, neural networks (also called Multi-Layer Perceptrons in this literature), radial basis function classifiers, and recently support vector machines.

3 Prior Research

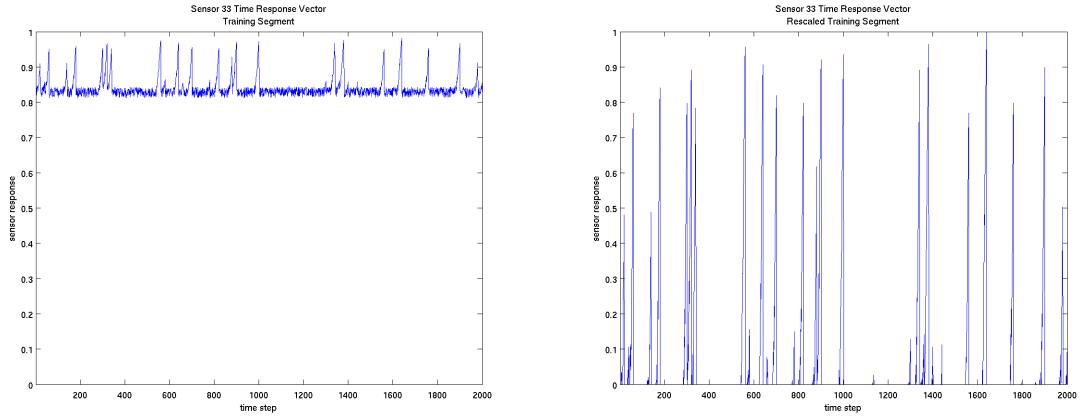
Many of the early research pieces on olfactory classification drew on the literature of artificial neural networks. [38], one of the early presentations of the olfactory sensing technology, recommends neural networks as showing promise for this application; examples include [21, 40, 45, 16]. Recent research also applies neural networks to this problem, including [44, 15, 4]. Other approaches include fuzzy logic [11, 27], nearest neighbors [33], fuzzy nearest neighbors [5], and genetic and evolutionary algorithms [10, 26]. Much recent research seems to be in the area of biologically inspired systems, perhaps a revisiting of the neural network approach [19, 39, 30, 32, 17, 35, 34, 28]. Finally, support vector machines have been successfully applied to this problem as they have to classification problems in other domains [8, 29].

4 Current Approach

4.1 Data Set

The data set studied in this project is a numerical simulation of results from a nanotechnology-based system as discussed in section 2, under development by Evolved Machines and partner iSense, for the DARPA RealNose program [1]. Real sensor data is not yet available, so a set of simulated data was made available for this project. The long-term objective would obviously be to develop and refine the classification algorithm on the synthetic dataset and apply it to real sensor data when the sensor array is ready.

The data set consists of separate testing and training data. In each data set, an array of 100 sensors is exposed to a sequence of 100 odorants. Each odorant is presented to the sensor individually in the presence



(a) Raw data

(b) After background subtraction

Fig. 1: Sensor 33 response over all training examples (time steps)

Algorithm 1 Background subtraction

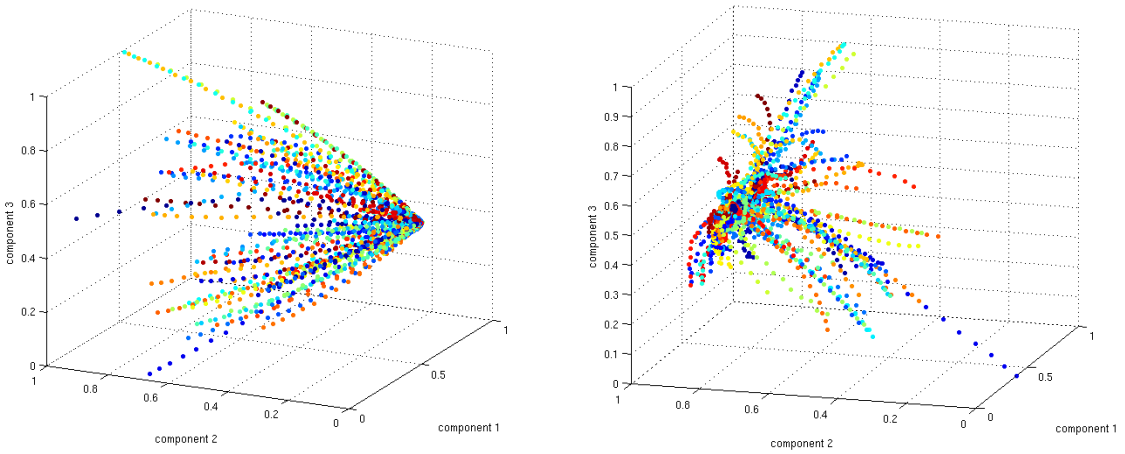
```
for each sensor
    for i=1 to maxBuckets
        group sensor data points in i buckets
        calculate differences between number of values in each bucket
        find pair of buckets a and b with largest difference
        calculate backgroundBoundary: upper boundary of bucket a
        calculate numBelow: number of data points below backgroundBoundary
        terminate when numBelow does not change for 5 iterations
    end
    discard data points below backgroudBoundary
    rescale remaining data points to [0, 1]
end
```

of a background of significant concentration consisting of two different, unknown odorants, and is presented in 20 log-increasing concentrations from 10^{-3} to 10^{-1}. The sensors are designed to mimic the response of human olfactory receptors; as a consequence, each sensor responds to a broad spectrum of odorants, so that the signature response of a particular odorant will be a combination of responses from multiple sensors. Additionally, the response of any particular sensor is non-linear, falling off as the concentration of an odorant to which it responds nears its saturation level. This data set, however, does not exhibit sensor drift; the real sensors on which this data set is modeled have been designed to eliminate this problem associated with earlier sensor technology. The data were also generated with sensor response values in the range [0,1], so rescaling was not necessary.

4.2 Classification Methods

The classification approach taken in this paper involves three steps: *preprocessing*, *feature selection*, and *classification*. As noted above, the target odorants are presented in the presence of a significant background, so before the sensor responses can be used in classification this background must be removed. See Figure 1 for an example of response data from one sensor before and after background subtraction. As can be seen from the figure, the background response varies within a relatively narrow range, so to remove the background we need to find the upper boundary of this range, which is done according to Algorithm 1.

To select relevant features from among the 100 sensor responses, we implemented principal component



(a) First three components of training data

(b) First three components of testing data

Fig. 2: Principal component analysis of testing and training data after background subtraction. Each like-colored ray of points represents a single olfactant at increasing concentrations.

analysis (PCA). We noted that while the training set would be assumed present in its entirety, and hence would be available for PCA decomposition, the training set is assumed to be presented one element at a time. Consequently, we approached PCA by finding the principal components of the training data, then generating the PCA values of the testing data using the principal components found from the training data. See figure 2 for a scatter plot of the first three principal components of the training and testing data. As is evident from the figure, PCA reveals some surprising structure to the data; many of the olfactants are clearly visually separable in the training data as arcs of equally-spaced points projecting from a single origin. This structure seemed to suggest that a SVM with an appropriate kernel and kernel parameters would easily separate the olfactants in PCA-space, but the results did not support this conjecture.

The classification algorithms implemented for this project are: a Naïve Bayes classifier (NB), multiclass support vector machines with linear (LSVM), sigmoid (SSVM), polynomial (PSVM) and radial basis function (RSVM) kernels, and for comparison a bagged decision tree (BDT). NB was implemented in Matlab. After significant experimentation with various platforms, including CVX [18], SVM^{light} [23], SVM^{struct}, and libSVM [6], the final SVM testing was done using the Matlab interface to libSVM. The bagged decision tree was implemented using the Matlab Statistics Toolbox. Results from each classifier are presented below.

Since a standard SVM is really a two-class classifier, several methods exist for using an SVM in a multi-class scenario. We investigated the two that appear most often in the literature, one-versus-all (OVA) and one versus one (OVO). In the OVA method, a classifier is trained on each class, using the training examples from that class as positive examples and all other training examples as negative examples. To classify a new example, each of these classifiers is run on the new example; the class of the new example is that of the classifier that returns the largest distance from the separating hyperplane. In the OVO method, also known as *max-wins-voting*, $\binom{n}{2}$ classifiers are trained, each of which takes one class as positive and another class as negative. To classify a new example, each classifier is run, and a win is counted for each class selected by its classifier the opposing class. The new example is classified as belonging to the class with the largest number of wins. There is significant literature on the question of which method is more accurate (see, for example, [24], [22], [13], [9], and [7]), but the results presented in these papers and particularly in the study by Rifkin and Klautau [36] suggest that relative to the selection of kernel and kernel parameters the multiclass classification method is insignificant. As a result, we used the OVO method, which is that implemented by libSVM.

Classifier	Test Accuracy	Optimal parameters	Best CV accuracy
NB	68.6%	N/A	N/A
LSVM	43.65%	$C = 2^7$	81.75%
PSVM (degree 2)	43.6%	$C = 2^{13}$ $\gamma = 2^{-7}$	82.0%
PSVM (degree 3)	42.6%	$C = 2^{13}$ $\gamma = 2^{-5}$	82.0%
PSVM (degree 4)	42.1%	$C = 2^{13}$ $\gamma = 2^{-5}$	82.0%
RSVM	42.9%	$C = 2^{13}$ $\gamma = 2^{-5}$	82.4%
SSVM	43.9%	$C = 2^{15}$ $\gamma = 2^{-7}$	79.95%
BDT	67.0%	50 trees	N/A

Tab. 1: Classifier results

4.3 Classifier Selection and Evaluation

Each classifier was trained on the 2000-sample training set. Additionally, for the SVM classifiers, extensive cross-validation (CV) analysis was done to search for an effective combination of parameters. Two CV methods were evaluated. In the first method, which we refer to as *leave-one-concentration-out cross validation* (LOCO-CV), twenty CV iterations were done, leaving out a different concentration level from each olfactant on each iteration. The classifier was trained each time on the remaining 1900 training examples and tested on the 100 examples left out. The second method was a standard randomized 5-fold cross-validation (R5-CV), in which the data was randomly divided into five groups, one of which was left out and used as the cross-validation set on each iteration. While these CV methods seem to offer substantially different methods of evaluating the classifiers, the results of both were substantially similar. As a consequence, for consistency the results presented below all use classifiers evaluated with R5-CV.

5 Results

The classification results for each of the tested classifiers are presented in Table 1. Although all the SVM classifiers demonstrated approximately 80% cross-validation accuracy, their generalization error was significantly higher than predicted by cross-validation. The boosted decision tree achieved good results, but the best results were achieved by the NB classifier.

The results presented above all use data after background subtraction, but without PCA decomposition. Despite the apparent structure of the data when separated into principal components, using data after PCA decomposition actually decreased the accuracy of all classifiers tested. Furthermore, we investigated using a geometric curve-fitting technique to identify the arcs visible in the PCA plot, but it is not clear how to translate the parameters for such an arc from the testing data space to the training data space.

6 Conclusion and Further Research

Previous work in this area, particularly that in [8], suggested that a SVM classifier should produce good classification results for olfactory sensor data. Our results did not confirm this hypothesis, but neither should they be understood as an invalidation of the hypothesis. Instead, these results should be understood as confirmation that the implementation and tuning of SVM classifiers poses significant technical and conceptual challenges, even when using a pre-existing SVM implementation, which were not fully surmounted during the course of this research project. The multitude of applications for an accurate and precise olfactory classifier clearly warrants further research into this area. Future iterations of this research project will attempt to understand more fully the shortcomings of the SVM classifiers implemented in this project, as well as investigating other background extraction and feature selection methods and classifiers.

7 Acknowledgements

This project would not have been possible without the coooperation and support of Paul Rhodes and Evolved Machines, who provided the problem description and data set as well as helpful advice throughout the project. Thanks are also due to Morgan Quigley, whose suggestions for research directions were invaluable at each stage of this research.

References

[1] Darpa realnose program. http://www.darpa.mil/dso/thrusts/bio/biologically/realnose/index.htm.

[2] S. Baker, G. G. Roberts, and M. C. Petty. Phthalocyanine langmuir blodgett-film gas detector. *IEEE Proceedings I: Solid State and Electron Devices*, 130(5):260–263, 1983.

[3] Jin Ho Bang, Sung H Lim, Erwin Park, and Kenneth S Suslick. Chemically responsive nanoporous pigments: colorimetric sensor arrays and the identification of aliphatic amines. *Langmuir : the ACS journal of surfaces and colloids*, 24(22):13168–72, 2008.

[4] S. Bashyal, G. K. Venayagamoorthy, and B. Paudel. Embedded neural network for fire classification using an array of gas sensors. In *Sensors Applications Symposium, 2008. SAS 2008. IEEE*, pages 146–148, Feb. 2008.

[5] Rossella Blatt, Andrea Bonarini, Elisa Calabró, Matteo Torre, Matteo Matteucci, and Ugo Pastorino. Fuzzy k-nn lung cancer identification by an electronic nose. In *Applications of Fuzzy Sets Theory*, volume 4578, page 261. 2007.

[6] Chih-Chung Chang and Chih-Jen Lin. *LIBSVM: a library for support vector machines*, 2001. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.

[7] Can Demirkesen and Hocine Cherifi. A comparison of multiclass svm methods for real world natural scenes. 5259:752, 2008.

[8] Cosimo Distante, Nicola Ancona, and Pietro Siciliano. Support vector machines for olfactory signals recognition. *Sensors and Actuators B: Chemical*, 88(1):30–39, 2003.

[9] Kai-bo Duan and S. Sathiya Keerthi. Which is the best multiclass svm method?: An empirical study. In *Proceedings of the Sixth International Workshop on Multiple Classifier Systems*, pages 278–285, 2005.

[10] D. Dumitrescu, B. Lazzerini, and F. Marcelloni. Olfactory signal classification based on evolutionary computation. In *IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339)*, page 313, 1999.

[11] D. Dumitrescu, B. Lazzerini, and F. Marcelloni. A fuzzy hierarchical classification system for olfactory signals. *Pattern Analysis & Applications*, 3(4):325, 2000.

[12] A. M. Fard. A study of semiconductor sensors for an electronic olfactory system. MSc Thesis, University of Warwick, 1985.

[13] V. Franc and V. Hlavac. Multi-class support vector machine. In *Object recognition supported by user interaction for service robots*, page 236, 2002.

[14] Walter Freeman. *Olfactory system: odorant detection and classification*, volume III. Academic Press, New York, 1999.

[15] Jun Fu, Guang Li, Yuqi Qin, and Walter J. Freeman. A pattern recognition method for electronic noses based on an olfactory neural network. *Sensors and Actuators B: Chemical*, 125(2):489, 2007.

[16] J W Gardner, E L Hines, and M Wilkinson. Application of artificial neural networks to an electronic olfactory system. *Measurement Science and Technology*, 1(5):446, 1999.

[17] Julian W. Gardner, James A. Covington, Su-Lim Tan, and Timothy C. Pearce. Towards an artificial olfactory mucosa for improved odour classification. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 463(2083):1713, 2007.

[18] Michael Grant, Stephen Boyd, and Yinyu Ye. Cvx: Matlab software for disciplined convex programming. http://cvxr.com/cvx, December 2009.

[19] Agustín Gutiérrez and Santiago Marco, editors. *Biologically Inspired Signal Processing for Chemical Sensing*, volume 188 of *Studies in Computational Intelligence*. Springer, 2009.

[20] R. Gutierrez-Osuna. Pattern analysis for machine olfaction: a review. *Sensors Journal, IEEE*, 2(3):189–202, Jun 2002.

[21] E. L. Hines, J. W. Gardner, and R. N. Stansfield. A standalone neural network based electronic nose. In *DSP (Digital Signal Processing) in Instrumentation, IEE Colloquium on (Digest No.009)*, pages 10–1, Jan 1992.

[22] Chih-Wei Hsu and Chih-Jen Lin. A comparison of methods for multiclass support vector machines. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 13(2):415–25, 2002.

[23] T. Joachims. Making large-scale svm learning practical. In B. Schölkopf, C. Burges, and A. Smola, editors, *Advances in Kernel Methods - Support Vector Learning*. MIT Press, 1999.

[24] Laine Kahsay, Friedhelm Schwenker, and Günther Palm. Comparison of multiclass svm decomposition schemes for visual object recognition. 3663:334, 2005.

[25] Tae Hyun Kim, Sang Hun Lee, Joohyung Lee, Hyun Seok Song, Eun Hae Oh, Tai Hyun Park, and Seunghun Hong. Single-carbon-atomic-resolution detection of odorant molecules using a human olfactory receptor-based bioelectronic nose. *Advanced Materials*, 21(1):91, 2009.

[26] Barry K Lavine, Charles E Davidson, Curt Breneman, and William Kaat. Genetic algorithms for classification of olfactory stimulants. *Methods Mol. Biol.*, 275:399–426, May 2004.

[27] B. Lazzerini, A. Maggiore, and F. Marcelloni. Fros: a fuzzy logic-based recogniser of olfactory signals. *Pattern Recognition*, 34(11):2215–2226, 2001.

[28] Xu Li, Guang Li, Le Wang, and Walter Freeman. Study of a bionic pattern classifier based on olfactory neural system. *Journal of Bionics Engineering*, (2):133–140, 2004.

[29] Amy Loutfi, Silvia Coradeschi, Achim j. Lilienthal, and Javier Gonzalez. Gas distribution mapping of multiple odour sources using a mobile robot. *Robotica*, 27(2):311–319, 2009.

[30] Mehmet K. Muezzinoglu, Alexander Vergara, Ramón Huerta, Thomas Nowotny, Nikolai Rulkov, Henry D. I. Abarbanel, Allen I. Selverston, and Mikhail I. Rabinovich. Artificial olfactory brain for mixture identification. In *NIPS*, pages 1121–1128, 2008.

[31] H.T. Nagle, R. Gutierrez-Osuna, and S.S. Schiffman. The how and why of electronic noses. *IEEE Spectrum*, 35(9):22, 1998.

[32] E. Neuhaus and H. Hatt. Mimicking the olfactory system for the classification of chemical data. *Trends in Biotechnology*, 26(7):347–349, July 2008.

[33] Carey E. Priebe. Advanced data analysis methods for analyte recognition from optical sensor arrays. 1998.

[34] Baranidharan Raman. *Sensor-based machine olfaction with neuromorphic models of the olfactory system*. PhD thesis, College Station, TX, USA, 2005. Adviser-Gutierrez-Osuna, Ricardo.

[35] Baranidharan Raman, Ping A Sun, Agustin Gutierrez-Galvez, and Ricardo Gutierrez-Osuna. Processing of chemical sensor arrays with a biologically inspired model of olfactory coding. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 17(4):1015–24, 2006.

[36] Ryan Rifkin and Aldebaro Klautau. In defense of one-vs-all classification. *J. Mach. Learn. Res.*, 5:101–141, 2004.

[37] H Shurmer, A Fard, J Barker, P Bartlett, G Dodd, and U Hayat. Development of an electronic nose. *Physics in Technology*, 18(4):170, 1987.

[38] H. V. Shurmer. An electronic nose: a sensitive and discriminating substitute for a mammalian olfactory system. *Circuits, Devices and Systems, IEE Proceedings G*, 137(3):197–204, Jun 1990.

[39] Zu Soh, T. Tusji, N. Takiguchi, and T. Tsuji. Neuro-based olfactory model for estimation of sensory characteristic of mice. In *2008 IEEE International Conference on Robotics and Biomimetics*, page 431, 2009.

[40] Jon M. Sutter and Peter C. Jurs. Neural network classification and quantification of organic vapors based on fluorescence data from a fiber-optic sensor array. *Analytical Chemistry*, 69(5):856, 1997.

[41] Victor V Sysoev, Bradly K Button, Kelly Wepsiec, Serghei Dmitriev, and Andrei Kolmakov. Toward the nanoscopic "electronic nose": hydrogen vs carbon monoxide discrimination with an array of individual metal oxide nano- and mesowire sensors. *Nano Lett.*, 6(8):1584–8, 2006.

[42] N. Taguchi. British patent no. 1 257 155, 1971.

[43] R. H. Tredgold, M. C. J. Young, P. Hodge, and A. Hoorfar. Gas sensors made from langmuir-blodgett films of porphyrins. *Solid-State and Electron Devices, IEE Proceedings I*, 132(3):151–156, June 1985.

[44] Le Wang, Guang Li, Xi Liu, Baojun Wang, and W J Freeman. Study of a chaotic olfactory neural network model and its applications on pattern classification. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 4:3640–3, 2005.

[45] Y Yao, W Freeman, B Burke, and Q Yang. Pattern recognition by a distributed neural network: An industrial application. *Neural Networks*, 4(1):103, 1991.

[46] Hyeonseok Yoon, Sang*Hun Lee, Oh*Seok Kwon, Hyun*Seok Song, Eun*Hae Oh, Tai*Hyun Park, and Jyongsik Jang. Polypyrrole nanotubes conjugated with human olfactory receptors: High-performance transducers for fet-type bioelectronic noses. *Angewandte Chemie*, 121(15):2793, 2009.