Cigarette Brand Identification Using Intelligent Electronic Noses

Dehan Luo
Dept of Mechanical and Electrical Engineering, School of Economical and Technology, University of Science and Technology of China, Hefei, Anhui, China
luodehan@ustc.edu.cn

H. Gholam Hosseini
Dept of Electrotechnology, Faculty of Science and Engineering, Auckland University of Technology,
Auckland, New Zealand
hgholamh@aut.ac.nz

John R. Stewart
School of Computer Science, Queen's University of Belfast, Belfast, Northern Ireland, UK
j.r.stewart@qub.ac.uk

Abstract
In this study we propose a practical approach to increase the performance of electronic noses (E-noses) in cigarette brand identification. A portable E-nose was employed to collect and classify aroma signals from different brands of cigarettes. Artificial neural networks (ANN) were employed and trained with raw data and extracted features from the data collected by the E-nose to identify the cigarettes. The Chinese cigarette industry is losing millions of dollars per year due to counterfeit cigarettes. Detecting illegal cigarettes in the field is difficult, but may be possible using portable E-noses. However, the differences between odours from counterfeit and genuine cigarettes are small and detection may prove difficult. This preliminary investigation succeeded in identifying four different types of cigarettes in the laboratory. The identification results obtained from the intelligent E-noise trained with an ANN using the extracted parameters were better than the ones obtained directly from the E-nose.

Keywords
Artificial Neural Network, Electronic Nose, Intelligent Instrumentation, Cigarette Brand Identification.

INTRODUCTION
The Chinese cigarette industry is currently suffering multi-million dollar annual lost revenues due to widespread production and sale of counterfeit cigarettes [1]. It would be a great advantage in combating this very widespread illegal activity if the authorities had access to simple, straightforward means of determining the legality or otherwise of suspect cigarettes and which could be used at the time and site of detection. It has proved to be almost impossible to identify counterfeit cigarettes either by visual inspection or by the aroma detectable by the human nose. Emissions of volatiles from cigarette tobacco comprise a wide range of chemical components making up a complex odour which makes discrimination between brands and/or counterfeit cigarettes difficult. One potential solution to the problem is the use of portable electronic nose devices.

While electronic nose technology has been available in bulky, laboratory-scale form for over twenty years, it is only since around 1999 that miniaturized, portable devices using various sensor technologies (conducting polymers, SAW, tin-oxide) have become available[2]-[4] and have gained renewed interest in both academic and industrial research areas [5,6]. There has been much recent research into the development of E-nose systems for odour detection and measurement [7]-[10].

In the human olfactory system there are around 10,000 sensors; these are non-selective but can be very sensitive to certain odours. Signals from these sensors, when they are exposed to a complex odour, are interpreted in the brain which identifies the characteristics of the odour. Recently, it has been shown that one odour sensor can recognise multiple odours but different odours are recognised by different combinations of odour sensors [11].

Electronic noses work in a similar manner to the human olfactory system. They usually consist of: an array of four to thirty two sensors which react in some repeatable way when exposed to an odour; a system (usually electrical) for polling and assembling the sensor responses; and, an associated computer program which interprets resulting signals. An e-nose system will typically comprise a number of different components including the sensors, pumps, valves, flow controller, air conditioner, control buttons and display panel. It will have software for monitoring the hardware, data pre-processing, statistic analysis, and other built in functions.

The tests described here were carried out using a Cyranose 320 electronic nose at the Queen's University of Belfast's QUESTOR environmental research centre. This device contains 32 carbon black/conducting polymer sensors. When the sensors are exposed to an odour the polymers swell to a greater or lesser extent, changing the physical separation of the particles of the carbon black surface coating, and hence their electrical resistance changes. The extent of such change depends on the materials used and the composition of the odours to which they are exposed. The E-nose will provide a characteristic "fingerprint" for odours arising from an individual chemical or from mixtures of chemicals. An E-nose built from broad response sensors can be trained or calibrated using characteristic samples of a potential odour source in air. When the E-nose is presented with an "unknown" odour some kind of pattern matching is used to determine if the fingerprint of the new odour matches one which is already known (i.e. on which the E-nose has already been trained). The recognition software should determine the best match with its library of known fingerprints and report the degree of confidence in the match. Provided the odour was caused by a known chemical, for which the sensor was appropriate, it should be possible to identify the chemical [12].

The human equivalent of the electronic nose is an olfactometry panel – a number of people selected for their ability to respond in a reasonably repeatable way when presented with various odour samples. Using humans in this way is subject to variations in the sense of smell between human beings and between individuals on different occasions, and for these reasons is not as objective as an electronic nose. However, in many cases humans are more sensitive and can detect odours at much lower concentrations than electronic noses. When artificial neural networks are trained with data from olfactometry panels there will be extra difficulties associated with the variations associated with human observers [13].

The application of portable E-noses in field conditions is more difficult than in a controlled laboratory environment. Variations in humidity, temperature and, especially, background odours and sample concentration mean that identification rate can be disappointingly low. It is, anyway, more difficult to correctly identify complex samples such as those from cigarettes which can be expected to show natural variation among batches and with age and storage conditions.

The aim of this work is to explore and improve the capability of a portable E-nose to identify different brands of cigarette. The sensor responses from the E-nose were analysed using artificial neural network (ANN) techniques. The data analysis were performed using MATLAB V6.5 and its neural network toolbox at the Electrotechnology Department, Auckland University of Technology (AUT), New Zealand. Data collected from the portable E-nose were used for feature extraction and principal component selection purposes. The extracted features and principal component selection were used to train the neural network providing an intelligent approach for identifying cigarettes [14]. The preliminary investigation described here was carried out on four brands of cigarettes under open laboratory conditions (no special atmospheric, humidity or temperature controls). We found that the rate of identification obtained using the neural network trained with extracted parameters was better than that obtained directly from the E-nose.

MATERIALS AND METHODS

Cigarette manipulation

The leaves of four different brands of cigarettes (with a total mass of about 10g each) were obtained from local suppliers and placed in different flasks. The flasks were closed tightly after introducing the tobacco and were held at room temperature (18 to 20 °C) for 6 hours before sampling.

The sample flasks were connected to the EDU inlet via a sample transfer line. The tobacco odours were collected onto the internal Tenax adsorbent tube, then desorbed to the Cyranose 320 E-nose via its detector line.

E-nose measurements

Figure 1 shows the experimental set-up for measuring and analysing cigarette odours using E-nose in the laboratory. The Cyranose 320 E-nose was coupled with an Airsense EDU pre-concentrator containing Tenax TA adsorbent.

The sample flasks were connected to the EDU inlet via a sample transfer line. The tobacco odours were collected onto the internal Tenax adsorbent tube, then desorbed to the Cyranose 320 E-nose via its detector line.

Figure 1 The E-nose and preconcentrator system in the laboratory.

The volatiles were thermally desorbed from the EDU and

were pumped to the sensor array in the Cyranose 320. All measurements were performed at 30°C. The Cyranose 320 E-nose was set up as shown in Table 1.

Table 1 Cyranose 320 parameters set up for sampling cigarette volatile from Tenax tube

E-nose parameters		Run time (Sec)	Pump speed
Baseline Purge time		100	M
Sampling Time	Draw1	50	M
	Draw2	50	H
Purge Time	1st Air intake	15	M
	2nd Air intake	47	H
	2nd Air intake	28	M
Digital Filtering		On	
Substrate Heater Temperature		On	30 °C
Training Repeat count		10	
Identify Repeat Count		1	
Identification Quality		medium	
Algorithm		Canonical	

In Table 1, M and H stand for 'medium' and 'high' pump speed respectively.

Data acquisition and analysis

The raw sensor response data were acquired and formed the "fingerprints" shown in Figures 2 and 3. The data were also saved into files for further analysis. The Cyranose 320 E-nose was trained with the sensor response data to obtain the patterns for each brand of cigarette which were then stored in its memory.

Figure 2 shows that the different brands of cigarette have very similar fingerprints – because the main odour components are similar. The result is that it would be difficult to discriminate them using the methods provided with E-nose software.

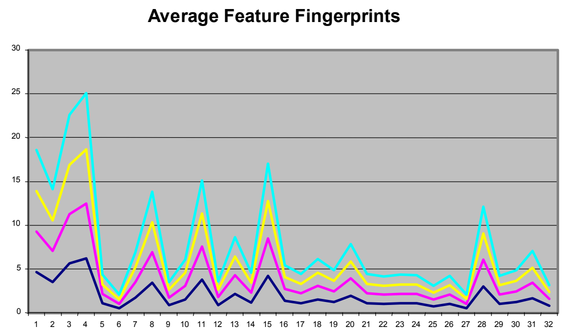

Figure 2 Fingerprints for four cigarette brands.

Figure 3 shows the average responses (from ten samples) from four brands of cigarette for each sensor. Note that all sensors, except S3, S4, S5, and S31, have very similar response data.

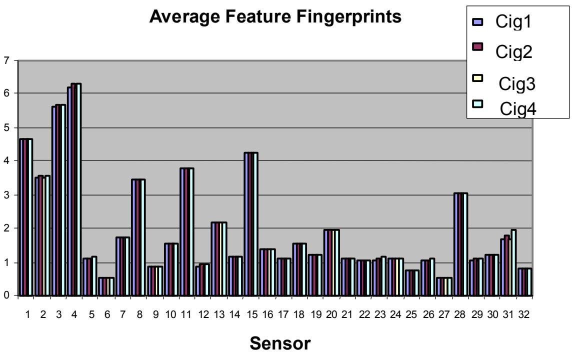

Figure 3 Average response for each brand.

Cigarette Identification using Cyranose 320 E-nose

The experimental tests for identifying the cigarettes were conducted in the laboratory one week after the Cyranose 320 E-nose had been trained. Identification samples were selected from the same packets of cigarettes as the training samples. Sampling time, desorption time and injection time were set to 100s, 80s and 60s respectively for EDU. The sampling time of Cyranose 320 was set to 100s. Each brand was sampled 10 times in order Cig1, Cig2, Cig3, and Cig4 with an interval of one day between test sets. The percentage correct identification rate is presented in Table 2. We can see that the E-nose canonical discrimination method has a poor discrimination rate, especially for Cig3 and Cig4 in this case.

Artificial Neural Networks (ANN)

In order to enhance the identification rate of cigarettes, a feed forward ANN has been used with the raw data from the E-nose. The nose has 32 sensors and collects 32 sets of data for each sample. These data were used to train our neural network for the purpose of comparing the identification rate. Several features were extracted from each sensor signal. They were the average (AVG), standard deviation (SD) and maximum response (MAX). As sensors S3, S4, S5, and S31 exhibited most resolution between the four brands, the output signals from these sensors were weighted for training purposes. All calculations and data evaluations were performed in MATLAB V.6.5 using the neural network toolbox.

The size of training data set and the number of extracted features was determined by the topology of the ANN. Figure 4 shows a two-layer log-sigmoid/log-sigmoid neural network. It has 32 or 35 inputs depending on using the additional three features with the 32 sensor signals. It has two neurons in its output layer to identify four classes. The number of neurons in its hidden layer was chosen to be eight. This number was selected by experience from other work.

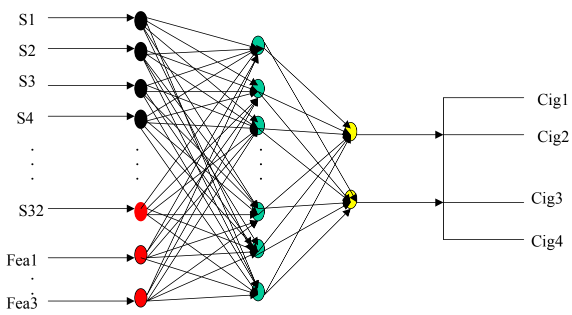

Figure 4 The network architecture with 32 inputs and 2 output neurons.

While altering the number of input neurons, the same transfer function and same parameters were used in the investigation. The log-sigmoid transfer function was chosen because its output range (0 to 1) is perfect for learning Boolean output values.

Four 32- or 35-element input vectors were defined as a matrix of input vectors for one sample to compare cigarette identification. The target vectors are two-element vectors to represent different classes. Each output vector corresponds to one class.

The same conditions were used to initialize networks with different architectures and the same parameters and learning functions were used to train them.

The Nguyen-Widrow initialization algorithm was used in this project. It has advantages over purely random weights and biases since few neurons are wasted (all the neurons are in the input space) and training works faster (each area of the input space has neurons). A log sigmoid transfer function in layers calculates a layer's output from its net input. Learning occurs according to the defined gradient, learning rate and momentum constant.

RESULTS AND DISCUSSION

Feature extraction and training of the ANN

The proposed networks were trained under the same conditions in order to compare their performances. The training data sets were extracted from the data files of Cyranose 320. Four brands of cigarette with 160 samples for each were selected as basic signals to train the networks. The first network had 32 input signals and the second network had extra inputs for three extracted features AVG, SD and MAX. The third network was the same as the second one but S3, S4, S5, and S31 signals were weighted.

Training parameters were chosen carefully with maximum epoch of 3000, minimum gradient limited to 0.00001 and goal of 0.0001. Figure 5 shows the training performances of the neural networks. We can observe that the behaviours of the three networks in terms of training speed are different.

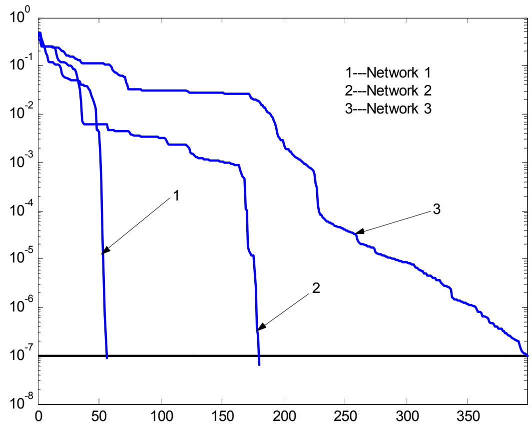

Figure 5 Training performance of the networks.

Test results

In comparing the performance of the ANN-based E-noses and Cyranose 320 E-nose, the same test data was used and each cigarette was tested 10 times. Each set of test data from the E-nose was preprocessed before submission to the ANNs for identification. Efficient features were extracted and some specific signals were weighted. The identification rate of each ANN-based E-nose is also shown in Table 2.

Table 2. Identification rates of cigarettes using Cyranose 320 and ANNs

Cyranose 320 E-nose	Cig1 (%)	Cig2 (%)	Cig3 (%)	Cig4 (%)
Without ANN	70	70	50	40
With ANN1	70	60	60	40
With ANN2	80	70	70	60
With ANN3	100	90	90	90

We can see the differences between results obtained by Cyranose 320 E-nose without and with three ANN networks. ANN1, with 32 input neurons, provides similar performance in correct identification of cigarettes as the Cyranose 320 E-nose without ANN. ANN2, which has 35 inputs (32 sensor signals plus three inputs for the extracted features) has better results than ANN1, but still its performance is rather low. ANN3, which uses weighted signals for S3, S4, S5 and S32 sensors, provides the best results. The identification rate has improved, especially for Cig1 and Cig4. Plainly, weighting the signals for specific sensors improves the identification rate.

Conclusions

In this work, we have demonstrated that artificial neural networks can improve E-nose classification performance for cigarette brand identification. In the experiments the performance of the Cyranose 320 E-nose was compared without and with utilising three ANNs in identifying four brands of cigarettes.

By utilising feature extraction techniques in combination with weighting sensitive signals from the E-nose and ANN, it is possible to improve E-nose identification performance. It is important to extract appropriate features from the raw data and apply them to the ANN for training and testing. The ANN input vector was formed by obtaining 32 sensor signals plus three features. Table 3 shows that there are different identification rates between ANNs with and without extracted features.

Significant improvement in successful discrimination between the brands was achieved by selecting specific sensors and applying extra weight to their signals during training and testing of the networks. The raw data from the Cyranose 320 indicated that signals from sensors S3, S4, S5 and S31 showed greater differences in their responses to the four different cigarette brands (than the other 28 sensors) and that S3 and S4 had the highest magnitude of response. Weighting these four signals was responsible for most of the advantage realised in the ANN tests.

The results from this study are sufficiently promising to justify further work with both legal and counterfeit cigarettes in China. The conditions used in the laboratory testing were not particularly controlled and the methods developed should be transferable to field analysis.

ACKNOWLEDGMENTS

Dehan Luo's visit to Queen's University and Auckland University of Technology (AUT) was funded by the China Scholarship Council. He acknowledges the support provided by the QUESTOR Centre at Queens and Electrotechnology Department at AUT. The E-nose used for the experimental work was supplied to the QUESTOR Centre by CyranoSciences Inc, USA.

REFERENCES

[1] State Tobacco Monopoly Administration (STMA) of China, 2001 Annual Report, *Proceedings of Annual*

Conference of Tobacco Supervision, Bejing, China, March, 2002.

[2] Lonergan, M. C., et al., "Array-Based Vapor Sensing Using Chemically Sensitive, Carbon Black-polymer Resistors," *Chem. Mater*, vol.8, 9, pp.2298-2312, 1996.

[3] Doleman, B. J., et al., "Quantitative Study of the Resolving Power of Arrays of Carbon Black –polymer Composites in Various Vapor Sensing Task," *Anal. Chem.* vol.70, 19, pp. 4177-4190, 1998.

[4] Natale, C. D., et al., "Electronic Nose and Sensorial Analysis: Comparison of Performances in Selected Case," *Sensors and Actuators*, B, vol. 50, pp. 246-252, 1998.

[5] Magan, N., et al., "Volatiles as an Indicator of Fungal Activity and Differentiation Between Species and the Potential Use of Electronic Nose Technology for Early Detection of Grain Spoilage," Journal of Stored Products Research, vol.36, pp. 319-340, 2000.

[6] Burl, M. C., et al., "Assessing the Ability to Predict Human Percepts of Odor Quality from the Detector Responses of a Conducting Polymer Composite-based Electronic Nose," *Sensors and Actuators*, B, vol.72, pp. 149-159, 2001.

[7] Misselbrook, T. H., et al., "Use of an Electronic Nose to Measure Odour Concentration Following Application of Cattle Slurry to Grassland," *Journal of Agricultural Engineering Research*, vol.66, pp. 213-220, 1997.

[8] Schaller, E., et al., "'Electronic Noses' and Their Application to food," *Lebensm.-Wiss. u.-Technol*., vol.31, pp. 305-316, 1998.

[9] Brezmes, J., et al, "Fruit Ripenness monitoring Using an Electronic Nose," *Sensors and Actuators*, B, vol. 69, pp. 223-229, 2000.

[10] Winquist, F., et al., "A Practical Use of Electronic Noses: Quality Estimation of Cod Fillet Bought Over the Counter," *Transducers '95 Eurosensors IX, The 8th International Conference on Solid-State Sensors and Actuators, and Eurosensors IX*, Stockholm, Sweden, vol. 1, pp. 695-697, June, 25-29, 1995.

[11] Doleman, B. J., et al., "Comparison of Odor Detection Thresholds and Odor Discriminablities of a Conducting Polymer Composite Electonic Nose Versus Mammalian Olfaction," *Sensors and Actuators*, B, vol.72, pp. 41-50, 2001.

[12] Cyrano Sciences Inc. USA, *The Cyranose 320 Electronic Nose User's Manual*, 3rd ed., (November 2000).

[13] Bockreis, A., et al., "Odour Monitoring by the Combination of Sensors and neural Network," *Environment Modeling & Software*, vol.14, pp. 421-426, 1999.

[14] Plaulsson, N., et al., "Extraction and Selection of Parameters for Evaluation of Breath Alcohol Measurement with an Electronic Nose," *Sensors and Actuators,* vol.84, pp. 187-197, 2000.