Application an Odour Recognition using Nearest Neighbour Algorithm

Anies Hannawati, Resmana Lim
Electrical Engineering Department, Faculty of Industrial Technology, Petra Christian University
e-mail: anies@petra.ac.id, resmana@petra.ac.id

Abstrak

Dalam tulisan ini dibahas sebuah proyek sistem pengenalan aroma menggunakan sensor hidung elektronik. Berbagai aroma diklasifikasi menggunakan sensor hidung elektronik jenis sensor gas oksida timah. Empat buah sensor digunakan untuk mengakusisi aroma yang selanjutnya diklasifikasi dengan algoritma nearest neighbour. *Sebagai metoda ekstraksi fitur digunakan* principal component analysis (PCA) *yang memungkinkan representasi data lebih kompak.*

***Kata kunci** : hidung elektronik, sensor gas oksida timah,* K-nearest neighbor, Principal Component Analysis

Abstract

This project was undertaken with the aim of developing control techniques and data processing methods to allow a relatively simple electronic nose developed for robotics applications to classify a wide range of odours. The electronic nose used tin oxide gas sensors as the main sensing elements. By modulating the sensor temperature inside the sensors, additional information was gathered which helps to identify unknown odours. A simple K-nearest neighbor algorithm was implemented as the basis of a pattern recognition system for recognizing odours presented to the system. Further, we considered ways of compressing the stored data and techniques for finding the best match among the trained data using Principal Component Analysis.

Keywords: electronic nose, tin oxide gas sensor, K-nearest neighbor, Principle Component Analysis.

Introduction

A small electronic nose was constructed and investigated to allow it to recognise different odours. An electronic nose commonly consists of a small number of sensors each with a distinct but broad and overlapping sensitivity to a range of chemicals. Though electronic nose systems are now commercially available for such applications as environmental monitoring, health care and quality assurance of foodstuffs, beverages and perfumes, the applications for these noses are mainly for industrial purposes and they are quite expensive (between US $20,000 and US $100,000 [1]). Only a few applications for electronic noses have been reported in robotic applications, such as mobile robots [2,3]. Rozas et al. [2] implemented an electronic nose on a mobile robot for locating the source of volatile chemicals.

For building electrical noses, many researchers have chosen commercially available tin oxide sensors [4]. In this case, the nose contained four tin oxide gas sensors which are Taguchi Gas Sensors (TGS) manufactured by Figaro Engineering Inc. The behaviour of tin oxide semiconductor is highly temperature dependent. This dependence was used to provide additional transient data. In an initial study, Russell [5] used this factor to investigate the possibility of building an electronic nose using a single tin oxide sensor.

In this project, sensor temperature is modulated which changes the characteristics of the sensors. This extra information allows many more chemicals to be discriminated and this data will form the basis for the recognition of unknown chemicals. Using this technique reduces the number of sensors required to distinguish a set number of odours and this in turn reduces the space, power consumption and expense of the electronic nose. The nose was trained on a number of household chemicals.

The rest of the paper is organized as follows. In section 2, experimental set up is explained. In section 3, we describe the algorithm used to recognize unknown odours. Experimental results are provided in section 4, and conclusions and future works are outlined in section 5.

Catatan : Diskusi untuk makalah ini diterima sebelum tanggal 1 Januari 2002. Diskusi yang layak muat akan diterbitkan pada Jurnal Teknik Elektro volume 2 nomor 1 Maret 2002

Experimental Set Up

1. Tin Oxide Sensor

Tin oxide sensors have a sensor resistance, whose value depends on the presence of oxygen [6]. The presence of oxygen increases the sensor resistance. As mentioned before, the behaviour of the sensors are highly temperature dependent. By varying the heater voltage of the sensors, the temperature of the sensing material is changed. Based on this idea, the transient response was established by changing the heater temperature through the full range from room temperature to normal operating temperature of the tin oxide sensors. During this interval, the sensor's responses were recorded. A preliminary experiment was conducted to determine the appropriate length of the heating and cooling periods. Figure 1 shows that the sensor took about 8 seconds to reach normal operating temperature after application of heater current. When cooling, it took about 20 seconds to return to room temperature.

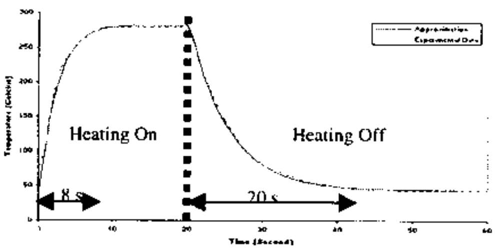

Figure 1. Thermal response for the sensor.

Based on the measured thermal response, the following cycle of the heating and cooling of the sensors was developed:

- The heater was switched on for 10 seconds (record sensor responses every 0.5 second). This was sufficient time for the sensor to be heated from room temperature to close to at normal operating temperature.
- The power was maintained to the heater and the fan was switched on for 5 seconds. This further period of heating helped the sensor to recover after exposure to a chemical.
- The heater was switched off and the fan was kept on for 16 seconds (to allow the sensor to cool).
- The heater and the fan were switched off for 4 seconds (to allow air movement to stop before the next sensing cycle). This gives a total cooling time of 20 seconds.

2. Electronic Nose

The electronic nose has four tin oxide gas sensors of types TGS 2600, TGS 2610, TGS 2611 and TGS 2620 manufactured by Figaro Engineering Inc. In addition, the nose contains a temperature sensor (LM35) manufactured by National Semiconductor Corporation and a humidity sensor (SMTRH 05), manufactured by LJK Technology. The body of the nose was formed by a metal tube coupled with a small electric fan. Figure 2. shows the cross section of the electronic nose.

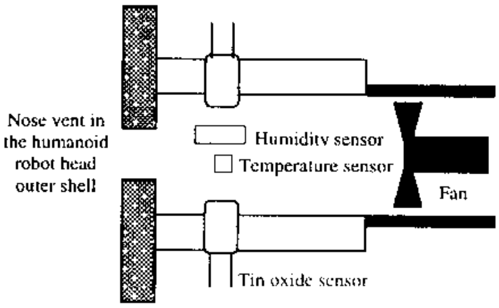

Figure 2. Cross section of the electronic nose.

The micro-controller used is MC68HC912B32. Digital outputs from Port A control power to the tin oxide sensor heaters and to the fan. The output of the four TGS sensors and the temperature sensor are fed into Port AD analogue to digital inputs of the microcontroller unit. The humidity sensor is a capacitive transducer, and this is read by using it as one of the timing components of a Schmitt oscillator circuit. The output of the Schmitt oscillator is fed into a timer port of the microcontroller unit. The block diagram of the hardware is shown in Figure 3.

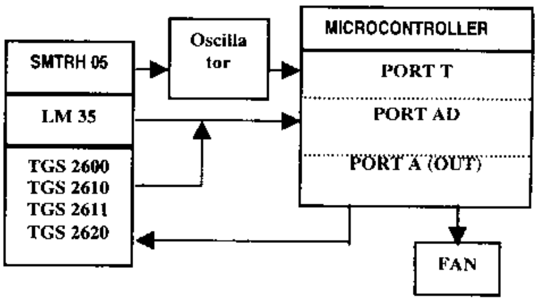

Figure 3. Block diagram of the electronic nose system.

3. Data Collection

In order to implement the electronic nose, a data base of stored patterns was established. The database contains information on the target household odours that the e-nose is to recognise. The cycle for sensor temperature developed previous section was used for gathering the stored patterns.

As mentioned before, during the heating period, 20 responses for each sensor were recorded. This process was repeated for a range of concentrations of each target chemical to build up a database of training data. Figure 4 shows one of the sets of training data obtained for one concentration of ethanol.

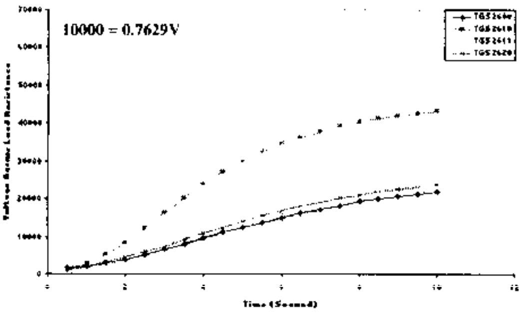

Figure 4. Responses of the sensors for one reading cycle.

Algorithms

1. Pattern Recognition System

A pattern recognition system is the basis of the software for recognising an unknown odour. Figure 5 shows the block diagram of this system.

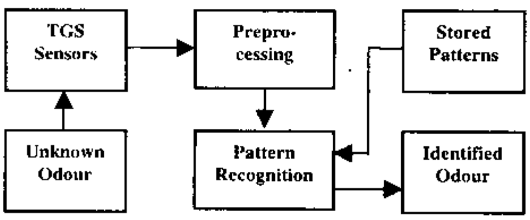

Figure 5. Pattern recognition system [3,8].

In this system an array of odour sensors, which convert volatile chemicals to electronic signals, is exposed to an unknown odour. The output of each sensor is then sampled by an analogue to digital converter so that the responses can be fed into a pattern recognition process. The pattern recognition process compares the pattern of sensor responses, produced by the unknown odour, with a database of stored patterns to determine the best match and thereby identify the unknown odour.

2. K-Nearest Neighbour

There are many pattern recognition techniques that could be used to match the sensor data produced by an unknown chemical with a database of training data. In electronic noses, the most common technique of pattern recognition is the artificial neural network [4]. However, the training process for this technique requires a lot of data and takes a long time. As a starting point for this project, a simple classification technique, the K-nearest neighbour algorithm, was used [3, 7].

This technique measures the distance between the unknown sample x_i and the stored pattern y_i. The Euclidean distance D between these two patterns is:

$$D = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \qquad (1)$$

where n is equal to 80 (20 readings from 4 sensors). The stored pattern, which has the smallest Euclidean distance from the unknown sample, is assumed to correspond to the same chemical with the same concentration as the unknown sample.

3. Principal Component Analysis

The method of modulating the sensor temperature has a disadvantage. It requires a search through a large database in order to recognise the unknown odour. Principal Component Analysis was used to address this problem.

To reduce the dimensionality of the feature vector, the first ten KLT (Karhunen Loeve Transform or Principal Component Analysis, PCA) coefficients of all training and test data feature were calculated. The KLT was used because it has the important property that the projection of the data set on the first *N* principle components has the highest energy concentration of any other N component projection. Therefore, it captures the highest amount of variation in a data set, more than any other linear transform for a fixed number of components.

To compute the KLT coefficients of the training feature, the first ten eigenodours were calculated using only the training data. Once the eigenodours were calculated, we projected the test and training feature onto the eigenodours. This projection produced ten coefficients for all the test and training smells features that when multiplied by their respective eigenodours and summed approximated the original test and training data. The ten coefficients for both the test and training data were used as inputs to the classification algorithms.

Results

For the experiment, training data was obtained from 13 different household chemicals that were acetone, ammonia, camphor, ethanol, coffee (ground coffee, espresso coffee, mocha coffee), tea (peppermint tea and cinnamon tea), oil (aromatherapy essential oil – lemon or yellow oil, room fragrance oil – red musk oil, eucalyptus oil), and incense.

The following figure is the first 8 of the principal components extracted from 1512 datasets (14 classes with 108 datasets each).

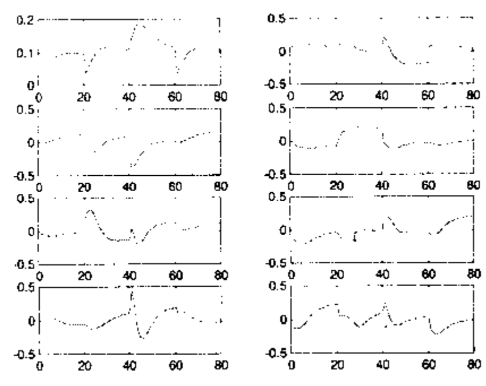

Figure 6. Principal Component Analysis Components.

The component number of eigenodours gives significant contribution to the pattern recognition results. Figure 8 shows that we got proper recognition results by using only between 5 and 10 components.

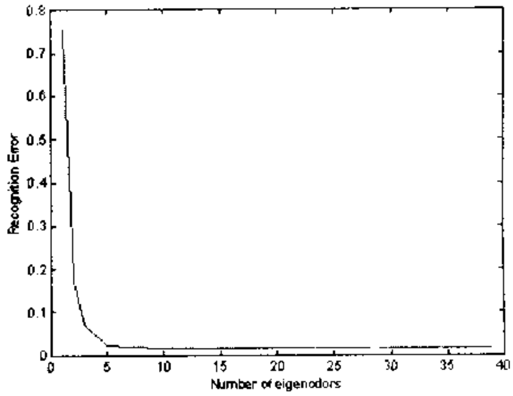

Figure 7. The effect of the eigenodours' component number.

In the experiment, the response resulting from exposure to each test chemical was recorded. These responses were compared with the stored patterns. The system successfully recognised the chemicals for a wide range of concentrations.

The next figure plots the features transformation into two dimensions of Principal Component Analysis. The class objects was clustered nicely thus they can be recognized easily.

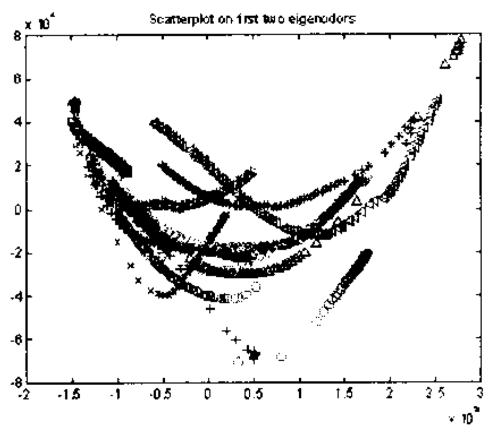

Figure 8. Two dimension of features transformation.

The electronic nose system that is being developed can recognize the trained chemicals with 92% recognition rate. For high concentration of chemical, the system successfully recognizes the unknown chemical. For low concentration of chemical, sometimes, we got incorrect results.

The K-Nearest Neighbour classification accuracy was 92.0% on the test set (95% on the training set).

Conclusions and Further Work

The electronic nose system developed is capable of identifying trained chemicals and that is able to distinguish their concentration over a wide range. The simple K-nearest neighbour method used works well for high concentrations of chemicals, but currently the results are less reliable for low concentrations of chemicals.

The method used in this project of recording the sensor response while changing the temperature of the TGS sensor has some advantages. This method allows many more chemicals to be discriminated and reduces the cost of the sensor system. It has been shown that the e-nose, even with a limited number of sensors, can discriminate between several target chemicals. However, this method has some disadvantages. It requires a search through a large database in order to recognise the unknown odour. The Principal Component Analysis method used can solve this problem and reduce the number of the stored patterns.

The data processing technique used was simple and easy to implement. In the future, a better algorithm for finding the best match among the trained data, such as an artificial neural network, should be explored to improve the performance of the e-nose. Hopefully, a better algorithm will provide more reliable results for low concentrations of chemicals.

References

[1] Kaplan, G., *The How and Why of Electronic Noses*, IEEE SPECTRUM, September 1998, pp 22-34.

[2] Rozas, R., Morales, J. and Vega, D., *Artificial Smell Detection for Robotic Navigation*, Fifth International Conference on Advanced Robotics, 1991, pp 1730-1733.

[3] Russell, R.A., Odour Detection by Mobile Robots, World Scientific, Singapore, 1999.

[4] Roger, K. and Erika, Handbook of Biosensors and Electronic Noses: Medicine, Food and the Environment, CRC Press, Boca Raton, 1997.

[5] Russell, R.A., *The World of Odour: a Relatively Unexplored Sensory Dimension for Robots*. 9th International Symposium or Robotics Research (ISSR '99), Snowbird, Utah, USA, 1999.

[6] Figaro Advertising Material, *General Information for TGS Sensor*, Http://www.figarosensor.com/products/general.pdf, 1997.

[7] Bwo, Sing-Tze., Pattern Recognition: Applications to Large Data-Set Problems, Marcel Dekker, New York, 1984.

[8] Gardner, J.W., and Bartlett, P.N., *Pattern Recognition in Odour Sensing*, Sensors and Sensory Systems for and Electronic Nose, Gardner, J.W., and Bartlett, P.N. (eds.), Kluwer Academic Publishers, 1992, pp.161-179.