Electronic Nose Testing for Confined Space Application Utilizes Principal Component Analysis and Support Vector Machine

M.A.A. Bakar, A.H. Abdullah, W.A. Mustafa, Z.B. Razali, S.A. Saidi, M.M.M.A. Kader, M.N.S.B.S. Aman
Department of Electrical Engineering Technology, Faculty of Engineering Technology, Universiti Malaysia Perlis (UniMAP), Kampus UniCITI Alam, Sungai Chuhuhu, 02100, Padang Besar, Perlis, Malaysia
E-mail: aizatbakar@unimap.edu.my

Abstract. A confined space has a limited space for entry and exit but it is large enough for workers to enter and perform work inside. It is not designed for continuous occupancy because it can contribute atmospheric hazards accidents that threaten the worker safety and industry progress. In this work, we reported the testing an instrument to assist workers for atmosphere testing during pre-entry. An electronic nose (e-nose) using specific sensor arrays is the integration between hardware and software that able to sense different concentrations of gases in an air sample using pattern recognition techniques. The instrument utilizes multivariate statistical analysis which is Principal Component Analysis (PCA) for discriminate the different concentrations of gases and the Support Vector Machine (SVM) to classify the acquired data from the air sample. The instrument was successfully tested using diesel, gasoline, petrol and thinner. The results show that the instrument able to discriminate an air sample using PCA with total variation for 99.94%, while the classifier success rate for SVM indicates at 98.21% for train performance and 95.83% for test performance. This will contribute significantly to acquiring a new and alternative method of using the instrument for monitoring the atmospheric hazards in confined space to ensure the safety of workers during work progress in a confined space.

1. Introduction
A confined space could contribute towards atmospheric hazard accidents because it has a limited means of entry or exit and is not designed for continuous occupancy but it is large enough for workers to enter and perform work [1]. Accidents that involve human fatalities or death sometime happen in these areas. The National Institute of Occupational Safety and Health (NIOSH) and Occupational Safety and Health Administration (OSHA) state that the safety of workers and environmental problem that threatens the industry progress effected from the presences of atmospheric hazards in confined space [2]. The atmospheric and physical hazards are included in two major types of hazards in confined space. Atmospheric hazards are more dangerous compared to physical hazards as they are unseen that presence from hazardous gases, oxygen deficiencies and dust. The cause of negative toxicological effects from hazardous gases can interfere with the human body’s ability to transport and utilize oxygen in human body. Normally the atmospheric hazards involve oxygen (O₂) are too low or high or the presence of flammable and toxic gases that workers are exposed to in confined spaces [3].
The methane (CH\textsubscript{4}) is the main flammable gas while the toxic gases include carbon monoxide (CO) and hydrogen sulphide (H\textsubscript{2}S).

Usually, the pre-entry for atmospheric testing is conducted before workers enter the confined space by carrying an instrument or device, so there is a need for a system that able to measure the hazardous gases. An electronic nose (e-nose) one of the example that able to measure and predicts the atmospheric hazards in confined space with high accuracy and repeatability [4]. Nowadays, variety applications of e-nose device can be found which include environmental monitoring, safety, plant disease detection and food quality assurance [5]. The e-nose functionality is able to mimic the discrimination of the mammalian olfactory system for smell is introduced in 1982 [6]. To sense the air sample in e-nose detection applications, a commercial set of gas sensors are used [7]. An e-nose carried by a mobile robot is the way on how technology can help to perform the pre-entry for atmosphere testing in confined space applications [8]. In this work, the experiment is to test the e-nose performance. Three objectives in this work which are to identify the capability of the sensor to react when exposed to different types of gas samples, to determine the ability of the e-nose in discriminating different types of gas samples and to establish the reliability of the e-nose that it can be trusted to be use in real environment for atmospheric hazards monitoring in confined space applications.

2. Electronic Nose System Structure
The e-nose device has been developed and tested at the Research Room II at School of Mechatronic Engineering, Universiti Malaysia Perlis (UniMAP) and Electrosoft Engineering, Sungai Petani, Kedah. The e-nose system structure includes the sensing module, signal conditioning, microcontroller and embedded software. Figure 1 shows the e-nose peripherals (i.e. purging system, air pump, power supply, input keypad, graphical Liquid Crystal Display (LCD) and wireless Radio Frequency (RF) communication are controlled by the microcontroller).

![Figure 1. E-nose system structure.](image)
Because of the sensors chamber need in porosity and inert characteristics, it is constructed by using polytetrafluoroethylene (PTFE) or Teflon material type. Aside, the sensors chamber is developed to ensure sensors stability, repeatability and reproducibility. The design also to ensure all sensors that placed inside it can be exposes to the air sample with an optimal sense. The gas sensors selection is measured with several criteria includes stability, sensitivity, response time, selectivity, recovery time, detection limit, life cycle and operating temperature [9]. The e-nose sensing module includes four selected Metal Oxide Semiconductor (MOS) gas sensors that are widely used in environmental monitoring [10]. The sensors selections as listed in Table 1 were based on the confined space main hazardous gases effective value with different sensitivity manufactured by Figaro Inc. and Synkera Technologies Inc. The sensors are able to detect main hazardous gases effective value to effect human for oxygen, carbon monoxide, hydrogen sulphide and methane at the 0-30%, 35 ppm, 20 ppm and 5%. Respectively.

| Category     | Target parameter | Sensor       | Sensitivity (ppm) | Sensitivity (%) |
|--------------|------------------|--------------|-------------------|-----------------|
| Oxygen       | Oxygen           | Sk-25F       | -                 | 0 to 30         |
| Toxic        | Carbon Monoxide  | TGS 2442     | 30 to 1000        | -               |
|              | Hydrogen Sulphide | PN 714       | 1 to 100          | -               |
| Flammable    | Methane          | TGS 2612     | -                 | 1 to 25         |

The e-nose control unit as system brain is using microcontroller (dsPIC33FJ128MC706A). For optimum response from the sensor, basically all the MOS gas sensor types must be heated to a certain temperature at a certain time [11]. To convert sensors response in terms of voltage signal into digital form to be acquired by the microcontroller, the analog to digital (ADC) interface is placed in between. The sensors response signals inform of voltage output will be converted by the microcontroller into 12 bit (4096) as ADC readings. The ZigBee (MRF24J40C) with 2.4 GHz IEEE Std. 802.15.4 ™ RF communication is used to send the readings via wireless communication to personal computer (PC). The data that received will be interpreted by control software in PC. Software Visual Basic 6.0 as Graphic User Interface (GUI) was designed and will help to present the readings in real time monitoring for more visualize.

3. Multivariate Statistical Analysis

3.1 Principal Component Analysis

The Principle Component Analysis (PCA) can be used to reduce a large set of variables to a small set that still contains most of the information in the large set and classified as a dimension-reduction tool. The PCA will loadings or weight of each original variable of variance percentage by transforming the number of original data with correlated variables (number of sensors) into uncorrelated variables (Principal Components, PC). The PCA is performed on a square symmetric matrix in traditionally. It can be a pure sum of squares and cross product (SSCP) matrix, covariance matrix (scaled sums of squares and cross products) or correlation matrix (sums of squares and cross products from standardized data). Since these objects only differ in a global scaling factor, the SSCP and covariance do not differ. If the variances of individual variants differ much or if the units of measurement of the individual variants differ, the correlation matrix will be used.

Basically the PCA can be represented in five steps which are normalize the data by subtracting the mean, calculate the covariance matrix, calculate the eigenvectors and eigenvalues of the covariance
matrix, choose components and form a feature vector and derive the new data set. The PCA result is being analysed and shown visually in two-dimensional (2D) or three-dimensional (3D) graphical plot that contains the original data’s important information. It also shows the group cluster and PC percentage. The PCA is the most widely used in e-nose testing for the dimension reduction technique.

3.2 Support Vector Machine

For data classification and regression, the Support Vector Machine (SVM) is a supervised learning model that is based on statistical learning [13]. For multi class data in separate group, the SVM training used data with known class when designing a linear classifier hyper plane model. Between hyper plane and closet training, the SVM is able to targeting clear maximum margins between both. Furthermore, in order to avoid over fitting, the SVM optimized classification decision function is based on structural risk minimization. For non-linear data cases, SVMs employ the kernel trick where a positive definite kernel function is used to map the input data into a high dimensional transformed feature space. The application is widely used in many fields since the SVM does not require any estimation of statistical distributions of classes to accomplish the classification task. This classifier performance is excellent in several applications for example determination of green tea quality grades [14] and identification of selected features on Mexican coffee [15].

4. Methodology

The e-nose performance testing experiment was conducted at the biomaterials laboratory, School of Mechatronics Engineering, Universiti Malaysia Perlis (UniMAP). The laboratory room temperature and humidity was continually observed and recorded maintain at 25°C with 75% Relative Humidity. Here is the data acquisition procedure to accomplish the experiment in data sampling process: -

i. Run the GUI on a laptop computer and switch on the e-nose to test the communication.

ii. Pre-heat the sensing module to ensure the MOS gas sensors reach the optimum operation temperature.

iii. Run the self-purging (purge cycle) by activating the 3-way solenoid electro-valve to allow the ambient air flow through the active carbon filter by using the air pump. The clear reference air will flow into the sensor chamber and flush out the previous air sample. Then the e-nose will remain idle for about 30 seconds to allow sensors response to return to its basic value.

iv. Run sniff cycle by turn off the 3-way solenoid electro-valve to switch channel for the e-nose expose to the air sample. By using the same air pump, the headspace air sample will flow into the sensor chamber.

v. The air sample will be stagnant in sensor chamber for 200 seconds that exposing to sensors for optimum interaction and sensor response reach steady state.

vi. The readings data is transfer wirelessly through the Zigbee module to a laptop computer. The GUI will record and display 200 data at one data per second baud rate based on the sensor’s response.

vii. Data will generate until 2000 data, repeat steps (iii) to (vi) for another nine times in order to be analyse in data processing part.

The main objectives and purpose of this experiment testing is to identify the capability of the sensor to react when exposed to different types of gas samples. Initially the ambient air sample is used as acquire reference data for the baseline manipulation. Then several gas sample were use includes diesel, petrol, gasoline and thinner that able to produces toxic and flammable gases as the sample. For each samples, a syringe was used to drip a 10 ml sample in the 100 ml beaker and covered with a plastic wrapper. To reach the liquid-vapour equilibrium stage in one beaker space, the sample was left in an idle state for 10 minutes. The beaker wrapper has one small hole for the e-nose nozzle to sniff the air sample from the headspace as shown in Figure 2.
5. Results and Discussion

Figure 3 shows the 2D PCA score plot for the system functionality testing. Four total PC’s that represent number of MOS gas sensors and first two PC’s variation indicates at 99.94% (PC1 scored 99.80% and PC2 scored 0.14%) which contain most of the useful information. The plot shows that the sample data successfully discriminated into four groups between diesel, gasoline, petrol and thinner. The diesel and petrol sample groups shows very closed because similarity of the characteristics and its applications. The plot indicates that the e-nose system functionality test was able to differentiate between different air samples.

To test the developed e-nose’s capability, the SVM as a statistical multivariate analysis method and robust features is suitable to use in analysing sample data. The analysis was conducted to classify the different air samples includes diesel (Die), gasoline (Gas), petrol (Pet) and thinner (Thin). The SVM method used 400 sample data inputs, 100 data per each samples. The test class was set at 30% (120 data) while train class was set at 70% (280 data). The SVM utilized the linear kernel function with prediction speed at 1900 obs/sec, training time at 21.918 sec, automatic kernel scale, box constraint at level one and the multiclass method was one versus one. Table 2 shows the system functionality test classification success rate is shown by the confusion matrix with 95.83% for test performance and 98.21% for train performance. The results indicate that the e-nose is successfully being tested with high classification success rate.
### Table 2. SVM confusion matrix for e-nose system functionality testing.

| Train Class | Test Performance |
|-------------|------------------|
| Die | 67 | 0 | 3 | 0 |
| Gas | 0 | 70 | 0 | 0 |
| Pet | 2 | 0 | 68 | 0 |
| Thin | 0 | 0 | 0 | 70 |

| Target Class | Test Class |
|--------------|------------|
| Die | 27 | 0 | 3 | 0 |
| Gas | 0 | 30 | 0 | 0 |
| Pet | 2 | 0 | 28 | 0 |
| Thin | 0 | 0 | 0 | 30 |

**Success rate**

Train: 98.21%

Test: 95.83%

### 6. Conclusion

In this work, the e-nose consists of four individual MOS gas sensors includes oxygen, carbon monoxide, hydrogen sulphide and methane have been successfully tested. The e-nose is capable of recognising the concentrations of several hazardous gasses by discriminated using PCA with a total variation of 99.94% while the classifiers success rate for SVM indicates of 98.21% for train performance and 95.83% for test performance. The results have proved the ability of the e-nose to be use in critical area like confined space during especially during pre-entry test. In the next stage, the e-nose will be integrated with a mobile robot for olfaction applications in a confined space and prove its reliability and functionality in real environment.

### 7. Acknowledgement

This work is partly funded under the Science Fund Research Grant Scheme (Project No: 03-01-15-SF0257) by the Ministry of Science, Technology & Innovation Malaysia (MOSTI). The authors also gratefully acknowledge to UniMAP for the opportunities given to do the project.

### References

[1] Workplace Safety and Health Council, Work. Saf. Heal. Counc. Collab. with Minist. Manpow. 6 (2010).

[2] A.J. Suruda, T.A. Pettit, G.P. Noonan, and R.M. Ronk, Hazard. Mater. 36, 45 (1994).

[3] H. Ye, in 2011 Int. Conf. Instrumentation, Meas. Comput. Commun. Control (2011), pp. 767–771.

[4] M. Mamat, S.A. Samad, and M.A. Hannan, Sensors 11, 6435 (2011).

[5] F. Rock, N. Barsan, and U. Weimar, Am. Chem. Soc. 108, 705 (2008).

[6] L. Capelli, S. Sironi, and R. Del Rosso, Sensors 14, 19979 (2014).

[7] K. Brudzewski, S. Osowski, and W. Pawlowski, Sensors Actuators B. Chem. 161, 528 (2012).

[8] K. Dandan, H. Albitar, A. Ananiev, and I. Kalaykov, Am. Sci. Res. J. Eng. Technology, Sci. 22, 210 (2016).

[9] G.H. Jain, in Fifth Int. Conf. Sens. Technol. (2011), pp. 66–72.

[10] K. Wetchakun, T. Samerjai, N. Tamaekong, C. Liewhiran, C. Siriwong, V. Kruefu, A. Wisitsoraat, A. Tuantranont, and S. Phanichphant, Sensors Actuators B. Chem. 160, 580 (2011).

[11] G.F. Fine, L.M. Cavanagh, A. Afonja, and R. Binions, Sensors 10, 5469 (2010).

[12] L.I. Smith, A Tutorial on Principal Components Analysis Introduction, 2002.

[13] C. Cortes and V. Vapnik, Mach. Learn. 20, 273 (1995).
[14] I.-C. Chen, J.K. Hill, R. Ohlemuller, D.B. Roy, and C.D. Thomas, Science (80- ). 333, 1024 (2011).

[15] R.B. Domínguez, L. Moreno-Barón, R. Muñoz, and J.M. Gutiérrez, Sensors (Switzerland) 14, 17770 (2014).