Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

An Investigation into the Suitability of Using Three Electronic Nose Instruments for the Detection and Discrimination of Bacteria Types

Geoffrey C. Green, Adrian D.C. Chan, and Rafik A. Goubran
Department of Systems and Computer Engineering, Carleton University, Ottawa, ON CANADA

Abstract—**The use of electronic nose (e-nose) technology for detection of food-borne bacteria has several practical advantages over current laboratory procedures, such as lower cost and reduced testing time. In this work, we are interested in using electronic nose systems to detect *E. coli* and *Listeria* in a nutrient broth, and discriminate between these bacteria types at various concentrations. To do this, we use instruments based on three different technologies - fingerprint mass spectrometry, metal oxide sensors, and conductive polymer sensors. Our results indicate that separation between groups can be achieved. We describe the relative merits and drawbacks of each technology and discuss how this rich multimodal dataset can be used to build a classification system.**

I. Introduction

ELECTRONIC nose (e-nose) technology is becoming an attractive option for analyzing many types of biological samples. At its heart, an e-nose system consists of: 1) a sample handling system; 2) an array of gas sensors; and 3) a pattern analysis and recognition (PARC) system. The sample handling system ensures that samples are presented to the sensors in a controlled manner. A wide range of gas sensor types exist in commercially available e-nose systems, including: conducting polymers, quartz crystal microbalance, surface acoustic waves, metal-oxide conductivity, and mass spectrometry. Each of these sensors have different selectivity patterns and yield a unique "odor signature" for the volatile organic compounds (VOCs) in the headspace of each sample under test. The PARC component is responsible for pre-processing sensor data, extracting features, dimensionality reduction, and classification decisions [1].

With an e-nose, analysis of samples of biological origin can be very rapid (typically minutes) and usually involves little sample preparation. In addition, a wide range of possible biological sample types can be analyzed. The potential to accurately diagnose medical conditions using data collected from an e-nose has garnered increasing research attention as of late [2]. While not yet in widespread clinical use, there are obvious potential advantages of this approach, including earlier detection and intervention by primary care physicians (and thus better health outcomes for patients), and eliminating the need for diagnostic lab tests (and the associated costs to the health care system). Efforts that have shown promise in this regard include determination of the presence of lung cancer from alveolar breath samples [3], detection of urinary tract infections from urine [4], and screening for diabetes with nose breath samples [5].

Though efforts at clinical diagnosis have dominated the e-nose research in the medical field, another potential application has arisen - the identification and classification of food-borne bacteria. Worldwide, 70% of bacterial infections are caused by food-borne bacteria, resulting in approximately 1.05 billion people being affected each year [6]. In Canada alone, there are about 10000 reported illnesses (the total number is certainly much higher) and 30 deaths attributed annually to the presence of unsafe organisms in food [7]. Of particular concern are the bacteria types *E. coli* (found in ground beef and raw fruits and vegetables) and *Listeria* (found in dairy products, vegetables fish, and meat). For detection and identification of these bacteria, traditional laboratory approaches (such as plating/culturing), while reliable and accurate, are also time-consuming and can be very expensive [8].

Initial attempts to use e-nose technology for bacteria identification have shown promise in a diverse range of applications. Dutta *et al.* reported considerable success (98% classification accuracy) when using a conducting polymer based e-nose to classify six bacteria types responsible for eye infections [9]. In another study, they used a similar technique on swab samples collected from patients' ENT (ear, nose and throat) regions in a hospital environment. The goal was to distinguish between two variants of *Staphylococcus aureus* with differing resistances to methicillin. Despite the fact that the samples still required culturing, this study exemplifies how e-nose testing has the potential to move closer to the patient and still achieve excellent results [10]. A couple of recent papers demonstrate the applicability of e-nose testing for food safety. Fend *et al.* performed a study that showed that *Myobacterium bovis* can be detected in cattle and badgers with e-nose analysis of blood serum samples [11]. Detecting the presence of salmonella in packaged meats (simulating a retail setting) was done by Balasubramanian *et al.* [12]. Using a relatively inexpensive conducting polymer–based instrument, they achieved classification accuracies nearing 90%. Finally, Alocilja *et al.* used an e-nose to differentiate pathogenic and non-toxic *E. coli* strains in different media [13].

In this paper, we assess the ability of three e-nose technologies to detect and discriminate between *E. coli* and *Listeria*. We compare the results of each and comment on our experiences with the different instruments. We investigated:

1. conductive polymer sensors (CPS) – the instrument used was the Cyranose 320 (Smiths Detection, New Jersey, USA,

formerly Cyrano Sciences) [14];
2. fingerprint mass spectrometry (FMS) – the instrument used was the AlphaMOS Kronos (AlphaMOS, Toulouse, France) [15]; and
3. metal oxide sensors (MOS) – the instrument used was the AlphaMOS FOX (AlphaMOS, Toulouse, France) [15].

The bacteria samples in this study are non-pathogenic and grown in nutrient broth. We report on our experiences with an eye towards future work involving pathogenic bacteria in food samples.

II. METHODS

A. Sample Preparation

In this study, non-pathogenic strains of *E. coli* (*E. coli DH5*) and *Listeria* (*Listeria innocua)* were provided by the Canadian Food Inspection Agency (CFIA), Ottawa, Canada. Both bacteria types were cultured at 37°C in Luria-Bertani (LB) broth containing 50mM MOPS (pH7.5) for 16-18 h and kept at 4°C prior to testing [19].

B. Electronic Nose Data Acquisition

Test 1: Detection and Discrimination of Bacteria Types

For each of the e-nose instruments, three sample replicates from the following classes (class label shown in parentheses) were tested: 1) *E. coli* in nutrient broth (E); 2) *Listeria* in nutrient broth (L); and 3) nutrient broth only (B) (control). For the MOS and FMS systems (which share the same sample handler – HS100, AlphaMOS, Toulouse, France), 1 mL aliquots were transferred by pipette into 10 mL vials which were then immediately sealed to prevent contamination. Each vial was agitated at 50°C/60°C (MOS/FMS) for 300s/900s (MOS/FMS) to release VOCs immediately prior to injection. On the MOS system, 1 mL of headspace was injected into the instrument at a rate of 1 mL/s (delivered to the sensors with pure air (O_2+N_2>99.95%, H_2O<5ppm, C_nH_m<5ppm, CO_2<5ppm) at a flow rate of 150 mL/min). Data were collected for 300 s (0.5 s sampling interval). On the FMS system, 4 mL of headspace was injected into the instrument at a rate of 0.1 mL/s (the carrier gas in this case was N_2). Mass fragments from 45-150 amu were scanned for 120 s. On the CPS system, sampling was performed at room temperature with no agitation. The inlet snout of the instrument was placed directly into the mouth of the flask containing 100 mL of sample), about 5 cm from the surface. After an initial purge of 20 s, the headspace was presented to the sensors and data collected for 45 s (the carrier gas in this case being room air), followed by a cleansing purge of 45 s. In all cases, the sample replicates were presented in alternating order (*i.e.* E, L, B, E, L, B, …).

Test 2: Effect of Sample Dilution on Detection and Discrimination of Bacteria Types

We were interested in investigating the extent to which varying the amount of bacteria in the sample affected our discrimination results. The original suspensions were diluted (by adding nutrient broth) to achieve concentrations of 10%, and 40% of the original concentration (by volume). To perform this test, we used only the FMS system with the same sample size, headspace generation, and data acquisition parameters, as described above.

C. Pre-processing

On the MOS and CPS systems, fractional difference responses were calculated to eliminate the effects of drifting baseline:

$$R_{preproc} = \frac{R - R_o}{R_o}$$

where R is the raw sensor reading, R_o is the baseline value, and $R_{preproc}$ is the value used in subsequent processing.

On the FMS system, no pre-processing was done.

D. Feature Extraction

The amount of raw data collected from each instrument makes it necessary to perform feature extraction to reduce the size of the data set for subsequent processing. On the MOS and CPS systems, the maximum absolute value of the (pre-processed) sensor response was chosen, yielding (for each sample) 12 features from the MOS system and 32 features from the CPS system. On the FMS, for each mass fragment, feature selection was done by finding the area under the intensity vs. time response curve over the time interval for which the intensity was half of its maximum or higher. This gave a total of 106 features. For all instruments, the features extracted were those suggested by the product's documentation or vendor consultation.

E. Dimensionality Reduction

All e-nose instruments use a large number of sensors – 12 (MOS), 32 (CPS), and 106 (FMS, considering each mass fragment reading as a sensor output) in this study – with differing selectivity, and this accounts for their ability to distinguish between wide ranges of odors. Unfortunately, this also causes problems during data analysis because the feature data are highly redundant and of high dimension. The difficulty of the subsequent classification task is exacerbated when a limited number of samples are used. To ensure that the full complexity of the classification boundaries is covered in a high dimensional feature space (and thus to achieve valid classification), the number of required training samples must grow exponentially with the dimension of the feature space. This dilemma is known as the *curse of dimensionality* in pattern recognition literature [16]. In this paper, feature reduction was performed using two stages:

1. Sensor Selection: On the MOS system, two sensors were faulty and thus removed from subsequent analysis. On the CPS system, all of the sensors were used. On the FMS system, those mass fragments whose responses were too noisy were eliminated – this was done by only selecting the

fragments with the highest intensity response (an approach taken based on vendor consultation) – in this study, we used the highest 15 responding fragments.

2. Principal Component Analysis (PCA): PCA is an unsupervised technique that performs a linear transformation on the feature vector. Eigenvectors of the feature space, with the highest eigenvalues, are identified, and the feature vector is projected onto those eigenvectors. Our analysis revealed that the first two or three most significant projections (called *principal components*) accounted for most of the variance (for all instruments) and these were retained.

III. RESULTS

Test 1: Detection and Discrimination of Bacteria Types

Figure 1 shows the result of performing PCA analysis on the three sample classes.

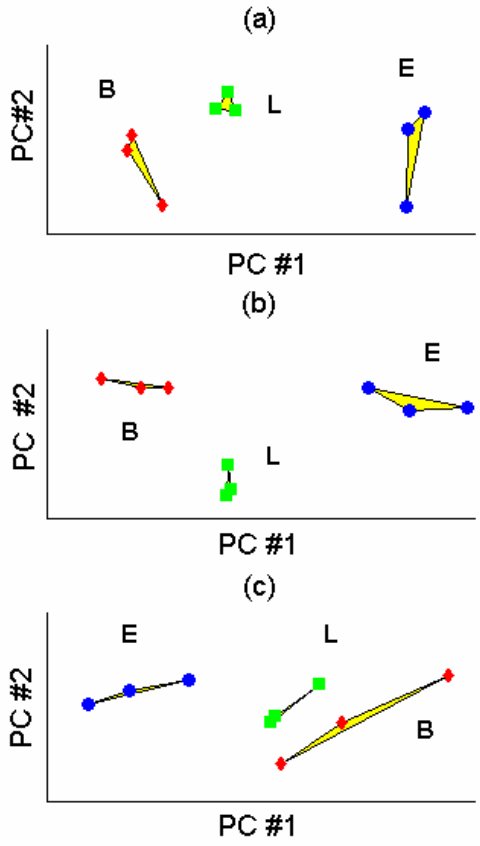

Figure 1: PCA plot (showing the first two principal components) for Test 1: a) MOS, b) FMS, and c) CPS.

Table 1 gives measurements of the separation between groups (using Euclidean distance between group centroids in the PCA space) as well as the repeatability of measurements (using standard deviation in the first principal component). These results indicate that differentiation between bacteria classes is possible with all three instruments. The group separation (relative to the size of each class cluster) is more pronounced for the MOS and FMS systems than it is for the CPS system.

Instrument	SD	Distance
FMS System		
E	0.93	4.08 (E-L)
L	0.04	3.12 (L-B)
B	0.62	5.20 (E-B)
MOS System		
E	0.0024	0.0437 (E-L)
L	0.0022	0.0208 (L-B)
B	0.0044	0.0636 (E-B)
CPS System		
E	0.0235	0.071 (E-L)
L	0.0127	0.032 (L-B)
B	0.0396	0.104 (E-B)

Table 1: Inter- and intra- group separation measures for Test 1.

Test 2: Effect of Diluting Samples on Detection and Discrimination of Bacteria Types

In this test, we grouped the data for the same bacteria strains together in the same class (*i.e.* there were 3 classes – E,L,B – and each had 9 samples, 3 from each concentration). We then performed PCA, giving the results in Figure 2.

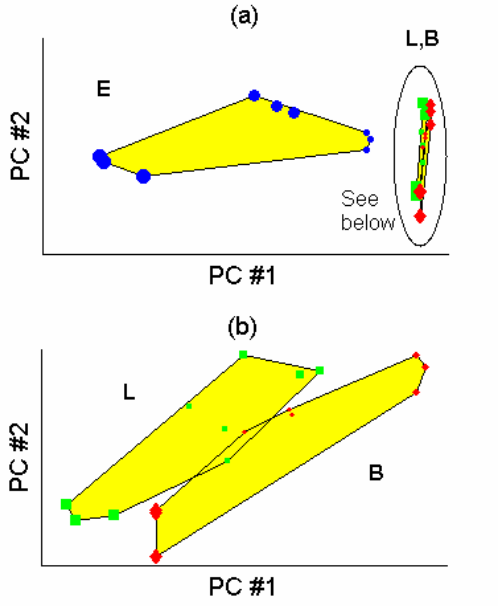

Figure 2: PCA plot (showing the first two principal components) for Test 2: a) all classes, b) zooming in on L and B classes. The size of the marker indicates the percentage of the original concentration used (large = original (100%), medium=40%, small=10%)

We can easily see that separation is quite clear between *E. coli* and the other classes, regardless of concentration. The difference between *Listeria* and nutrient broth is not as distinct; although subsequent analysis performed using three principal components (results not shown) indicate that these classes can in fact be linearly separated. It is also interesting

to note that in all classes, discrimination based on concentration is possible (note marker sizes in Figure 2).

IV. DISCUSSION

This study constitutes an initial investigation of the suitability of using three different e-nose technologies for the detection and discrimination of the bacteria *E. coli* and *Listeria* grown in nutrient medium. Our results indicate that both the detection of bacteria as well as separation between bacteria classes (including of varying concentration) is viable using all three technologies, a result consistent with other studies. A logical next step for this work would be to use the data collected to construct a classifier (such as those based on multilayer perceptrons, radial basis functions, nearest neighbours, or self-organizing maps [1]). Unknown samples could be projected onto these models (which would clearly require much more training data) to evaluate the classification accuracies possible with each. E-nose analysis usually shows that a small number of sensors have the most discrimination power for a given application [16]. With the outputs of all three technologies, a rich set of sensor data is available to us that could be combined to create a hybrid classifier with better performance than either instrument acting on its own. To this end, the technique of multi-dimensional combining used by Chen *et al*. is particularly promising [17]. It should be emphasized that the PCA performed in this paper is entirely *unsupervised* - the introduction of supervised methods (*e.g.* Fisher's linear discriminant analysis (LDA) or uncorrelated LDA (ULDA) [16,18]) would surely lead to better class separation.

Our experiences with these technologies have helped us understand the relative merits and disadvantages of each. In terms of sampling time, the FMS and CPS systems gave very fast results (~5 min), while the MOS system took longer (~20 min) because of the time required to return to sensor baseline. The FMS instrument is based on mass spectrometry, and has the potential ability to identify the chemical origin of the mass fragments in the headspace. This is not possible with either MOS or CPS systems. The robotic autosampler used by the MOS and FMS instruments allows for precise settings of both headspace generation and injections, but such control was not easily achieved with the CPS instrument. Though the MOS and FMS systems gave the best results in this feasibility study, further refinements in testing with the CPS (*e.g.* optimal headspace generation and sensor selection) would likely lead to better outcomes. The variations introduced by manual sampling likely account for the inferior results achieved with this method. Initial testing with the CPS system was attempted using a sample size of 1 mL in a 10 mL (as was used with the MOS/FMS instruments). Without the same combination of controlled sampling, agitation, and heating, we found that the headspace generated was not strong enough to give a reliable result. Finally, the handheld CPS instrument was the least expensive (and most portable) of the three, and the most expensive was the FMS.

Finally, it is noted that if e-nose technology is ever to be widely deployed for bacteria detection in food, sample preparation must be simple and quick. In future work, actual infected food samples should be used (instead of bacteria grown in nutrient) to better simulate real-life safety testing.

ACKNOWLEDGMENT

We would like thank Dr. Steven Luo and Dr. Min Lin from the Canadian Food Inspection Agency (CFIA) for providing the bacteria samples. This research was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Canada Foundation for Innovation (CFI), and the Ontario Innovation Trust (OIT).

REFERENCES

[1] H.T. Nagle, R. Gutierrez-Osuna, S.S. Schiffman, "The how and why of electronic noses," *IEEE Spectrum*, vol. 35, pp. 22-31, Sept. 1998.

[2] A.K. Pavlou, A.P. Turner, "Sniffing out the truth: Clinical diagnosis using the electronic nose," *Clin. Chem. Lab. Med.*, vol. 38, pp. 99-112, February 2000.

[3] C. Di Natale et al., "Lung cancer identification by the analysis of breath by means of an array of non-selective gas sensors," *Biosens. Bioelectron.*, vol. 18, pp. 1209-18, September 2003.

[4] A. K. Pavlou et al., "Use of an electronic nose system for diagnoses of urinary tract infections," *Biosens .Bioelect.*, vol.17, pp. 893-9, 2002.

[5] W. Ping, T. Yi, X. Haibao, S. Farong, "A novel method for diabetes diagnosis based on electronic nose," *Biosens. Bioelectron.*, vol. 12, pp. 1031-6, 1997.

[6] A. K. Deisingh and M. Thompson, "Biosensors for the detection of bacteria," *Can. J. Microbiol.*, vol. 50, pp. 69-77, 2004.

[7] http://www.inspection.gc.ca/english/corpaffr/foodfacts/ecolie.shtml

[8] A. K. Deisingh and M. Thompson, "Detection of infectious and toxigenic bacteria," *Analyst*, vol. 127, pp. 567-81, 2002.

[9] R. Dutta, E. L. Hines, J. W. Gardner, and P.Boilot, "Bacteria classification using Cyranose 320 electronic nose," *Biomed. Eng. Online*, vol. 1, pp.1-7, 2002.

[10] R. Dutta, D. Morgan, N. Baker, J. W. Gardner, and E. L. Hines, "Identification of *Staphylococcus aureus* infections in hospital environment: electronic nose based approach," *Sensors and Actuators B*, vol. 109, pp. 355-62, 2005.

[11] R. Fend et al., "Use of an electronic nose to diagnose *Myobacterium bovis* infection in badgers and cattle", *Journal of Clinical Microbiology*, vol. 43, pp. 1745-51, 2005.

[12] S. Balasubramanian, S. Panigrahi, C. M. Logue, M. Marchello, and J. S. Sherwood, "Identification of *Salmonella*-inoculated beef using a portable electronic nose system," *J. Rapid Methods Auto. Microbiol.*, vol. 13, pp. 71-95, 2005.

[13] E. C. Alocilja, N. L. Ritchie, and D. L. Grooms, "Protocol development using an electronic nose for differentiating *E. coli* strains," *IEEE Sensors Journal*, vol. 3, pp. 801-5, December 2003.

[14] http://www.smithsdetection.com

[15] http://www.alpha-mos.com

[16] R. Gutierrez-Osuna, "Pattern Analysis for Machine Olfaction," *IEEE Sensors Journal*, vol. 2, pp. 189-202, June 2002.

[17] H. Chen, R. A. Goubran, and T. Mussivand, "Improving the classification accuracy in electronic noses using multi-dimensional combining", in *Proc. 3rd IEEE Conf. On Sensors*, Vienna, Austria, 2004, pp. 587-90.

[18] J. Ye, T. Li, T. Xiong, and R. Janardan, "Using uncorrelated discriminant analysis for tissue classification with gene expression data", *IEEE/ACM Trans. Computat. Biol and Bioinform.*, vol. 1, pp. 181-190, 2004.

[19] M. Lin et al., "Monoclonal antibodies binding to the cell surface of *Listeria moncytogenes* serotype 4b", *Journal of Medical Microbiology*, vol. 55, pp. 291-9, 2006.