A Survey On Various Signal Pre-Processing Methods For Sensor Response In E-Nose

R. Sheebha[1] and Dr. M. Subadra[2]

[1.] Department of Electronics and communication Engineering, Einstein college of engineering Tirunelveli
Email: sheebha.queen@gmail.com
[2.]Head of the Department, Department of Electronics and communication Einstein college of engineering Tirunelveli
Email: subadra_m@yahoo.com

Abstract: This paper explores signal response and methods for extracting the desired digital signal, from gas sensor arrays, while maintaining the shape and resolution of thatsignal. The choice of signal preprocessing is critical and can have a significant impact on the performance of subsequent modules in the pattern analysis system. A comparative evaluation of Baseline manipulation, Compression and Normalization techniques for removing noise from signals during the pre-processing phase is provided. In E-Nose technology signal preprocessing is to extract relevant information from the sensor responses and prepare the data for multivariate pattern analysis.

[R. Sheebha, M. Subadra. **A Survey On Various Signal Pre-Processing Methods For Sensor Response In E-Nose.** *Researcher* 2014;6(11):76-80]. (ISSN: 1553-9865). http://www.sciencepub.net/researcher. 12

Keywords : E-Nose,baseline, sensor response, Stability.

I. Introduction

The response of e-nose sensors to odorants is generally regarded as a first order time response. The first stage in odor analysis is to flush a reference gas through the sensor to obtain a baseline [1]. The sensor is exposed to the odorant, which causes changes in its output signal until the sensor reaches steady-state. The odorant is finally flushed out of the sensor using the reference gas and the sensor returns back to its baseline. The time during which the sensor is exposed to the odorant is referred to as the response time while the time it takes the sensor to return to its baseline resistance is called the recovery time [2].

The next stage in analyzing the odor is sensor response manipulation with respect to the baseline. This process compensates for noise, drift and also for inherently large or small signals (Pearce et al., 2003). The three most commonly used methods as defined by Pearce et al. (2003) are as follows. Typically, raw signals acquired from gas sensors are contaminated by noise and outliers and as a result the signal is occluded to a significant degree making accurate measurement of a sensor's response impossible[18].

Noise in sensor systems has several possible sources and is introduced at various stages in the measurement process. Several forms of noise, including thermal and shot noise, are irreducible because they are inherent to the underlying physics of the sensors or electronic components. Other forms of noise which could be avoided originate from processes, and include 1/f noise, transmission, and quantization noise[1].

Noise introduced in the early measurement stages is considered to be the most harmful as it propagates and can be amplified through subsequent stages in the signal pathway[2].

Several signal processing approaches have been investigated as an approach to reducing noise levels [3].However, these approaches are typically static or steady state approaches and therefore do not encompass the full temporal signal [4].

In this paper we report on an evaluation of methods for feature extraction and denoising the digital signal from gas sensor devices.

Although signal preprocessing is somewhat dependent on the underlying sensor technology, three general stages can be identified: baseline manipulation, compression, and normalization.

The e-nose system is designed so that the overall response pattern from the array is unique for a given odor in an family of odors to be considered by the system.

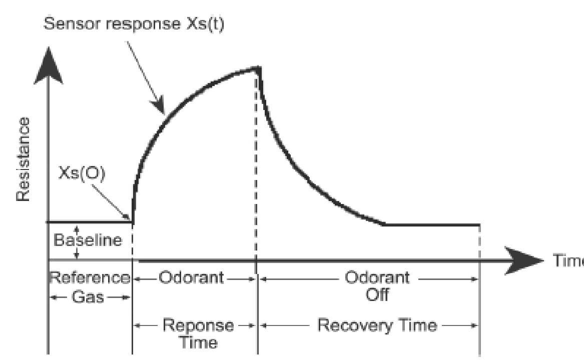

Fig: 1 E-nose sensor response to an odorant

The main objective of this work is to analyse and differentiate the various pre-processing of signal. Different type of pre-processing techniques are given, followed by the comparison among the methods to provide a comprehensive review.

II. Literature Survey

Enobong Bassey, Jacqueline Whalley and Philip Sallis in paper titled "An Evaluation of Smoothing Filters for Gas Sensor Signal Cleaning"[1] proposed the comparative evaluation of Savitzky–Golay smoothing, moving average, local regression and robust local regression filters for cleaning signals obtained from gas sensor devices during the pre-processing phase is provided. It was found that the Savitzky–Golay smoothing filtering method provided the best approximation of the sensor response.

In paper titled Stability Analysis of Metal Oxide Gas Sensors Using System Identification by Nimisha Dutta & et al.implementediterative prediction-error minimization (PEM) method an automatic pixel-based classification method for detecting unhealthy regions in leaf images. This proposed system is composed of state space model and calculating the transfer function. The model estimation is done by the system identification technique and the stable transfer function is determined based on the pole zero plot and the overshoot percentage.The work focuses on the derivation of various transfer functions under stable and unstable modes by PEM technique [2].

In [4], authors proposed three feature extraction methods of sensors, extraction from original response curves of sensors, curve fitting parameters and transform domains for pathogen detection based on electronic nose. By using the integrals, coefficients of exponential fitting with two parameters, hyperbolic tangent fitting, Fourier coefficient and Wavelet coefficient as features 100% identification accuracy was reached by using radical basis function neural network classifier[4].

In [5] "An Electronic Nose for Reliable Measurement and Correct Classification of Beverages" proposed the discriminative ability of the E-nose was evaluated using Principal Component Analysis and a Multi-Layer Perception Neural Network, with both methods showing good classification results. They have done relative baseline manipulation for sensor response.

In [6], authors developed a Fast and accurate method for E-nose system. They have developed Normalization for pre-processing the sensor response. In that sensor auto scaling can be performed in order to improve the pattern analysis. The classifier relies on a multilayer neural network based on a back propagation algorithm with one hidden layer of four neurons and eight neurons at the input and five neurons at the output.

Ricardo Gutierrez - Osunadescribes considerable number of methods from statistical pattern recognition, neural networks, chemo metrics, machine learning, and biological cybernetics has been used to process electronic nose data.The pre-processing is having three general steps can be identified as baseline manipulation, compression, and normalization was performed in this paper. Three baseline manipulation methods are commonly employed: difference, relative, and fractional can be used to eliminate drift from the sensor response [7].

Dongmin Guo and others has proposed, Baseline manipulation and normalization. Baseline manipulation is implemented for drift compensation, contrast enhancement, and scaling. Normalization is used to compensate for sample-to-sample variations caused by anolyte concentration. In this paper, we employed principal components analysis (PCA) to extract characteristic features of samples. Although the current pattern recognition method produced satisfactory results when they used integral data, it should still be possible to further improve the classification accuracy and speed by selecting proper features[8].

III. Types of Signal pre-Processing Techniques

A. Baseline Manipulation

The first stage of preprocessing consists of manipulating the sensor response with respect to its baseline (e.g., response to a reference analyte) for the purposes of drift compensation, contrast enhancement and scaling. Three baseline manipulation methods are commonly employed: difference, relative, and fractional. The *difference* method directly subtracts the baseline and can be used to eliminate additive drift from the sensor response. *Relative* manipulation, on the other hand, divides by the baseline, removing multiplicative drift, and generating a dimensionless response. *Fractional* manipulation, finally, subtracts and divides by the baseline, generating dimensionless and normalized responses [9,18].

Relative and fractional methods are useful to compensate temperature on the sensors, and the fractional method can linearize relationship between the resistance of the metal oxide sensor and the odor concentrations[4]. This method provides the good recognition results for the discrimination of several types of odors. Another method is log difference method. Log difference method is suitable when the variation of concentration of the odor-producing material is very large because it is able to linearize the highly nonlinear relationship between the odor concentration and sensor output[,20,21].

Method	Formula
Difference	$X_{ij} = V_{ij}^{max} - V_{ij}^{min}$
Relative difference	$X_{ij} = V_{ij}^{max} / V_{ij}^{min}$
Fractional difference	$X_{ij} = (V_{ij}^{max} - V_{ij}^{min})/V_{ij}^{min}$
Log difference	$Log(V_{ij}^{max} / V_{ij}^{min})$
Normalization	$X_{ij} = (V_{ij} - V_{ij}^{min})/(V_{ij}^{max} - V_{ij}^{min})$

The choice of baseline manipulation technique and the response parameter (e.g., resistance, conductance, frequency) is highly dependent on the sensor technology and the particular application. Fractional methods for MOS chemo resistors are also widely used. In the case of conducting polymerchemo resistors, fractional changes in resistance are commonly employed, both in research prototypes and in commercial instruments. Differential measurements are also widely used for MOSFETs [16,17,18].

B. Compression

The second stage in preprocessing is aimed at compressing the sensor-array response down to a few descriptors to form a feature vector or fingerprint. In most cases this is performed by extracting a single parameter (e.g., steady-state, final, or maximum response) from each sensor, disregarding the initial transient response, which may beaffected by the fluid dynamics of the odor delivery system[14]. However, with careful instrument design and sampling procedures, transient analysis cansignificantly improve the performance of gas sensor arrays: Improved selectivity: The dynamic response to an odor exposure (and the subsequentodor recovery) carries a wealth of odor-discriminatory information that cannot always be captured with a single parameter[15].

In some situations, transient parameters have also been reported to exhibit better repeatability than static descriptors. Therefore, sensor transients can be used as dynamic fingerprints to improves electivity by pattern-recognition means[9,10,11]. Reduced acquisition time: The duration of the acquisition cycles can be significantly shortened if the initial sensor transients contain sufficient discriminatory information, avoiding the lengthy acquisition times required to reach steady state. As a consequence, the sensors also require less time to recover their baseline, aprocess that can be particularly slow when the target odors have high concentrations. Increased sensor lifetime. By reducing the duration of the odor pulse and, therefore minimizing irreversible binding, the lifetime of the sensors can also be increased[18,19].

Various compression algorithms can be employed to generate descriptive parameters from the sensors' transient response. The standard procedure is to select the *steady-state* response of the sensor, but a number of compression algorithms have been proposed to extract additional information from the *transient response*, resulting in improved selectivity, reduced acquisition time, and increased sensor lifetime[15].

According to the procedure employed to generate the dynamic fingerprint, transient compression methods can be broadly grouped into three classes: Sub-sampling methods, Parameter-extraction methods and System-identification methods[17].

Sub-sampling methods exploit dynamic information by sampling the sensor transient response (and/or its derivatives) at different times during the odor exposure and/or odor recovery phase. *Parameter-extraction* methods compress the transient response using a number of descriptors, such as rise times, maximum/minimum responsesand slopes, and curve integrals. *System-identification*methods fit a theoretical model (e.g., multiexponential,auto-regressive) to the experimental transients and use the model parameters as features[18,19].

Exponential curve-fitting methods can result in nearly lossless compression of thesensor transients, but are computationally intensive[18]. For these reasons, subsamplingand parameter-extraction methods are more commonly employed. A final word of caution regarding the use of transient information: a large number of dynamicparameters will require an exponentially increasing number of training examples inorder to prevent the pattern recognition system from over-fitting the data. Alternatively, one may use resampling techniques (e.g., cross-validation, bootstrap) or regularization(e.g., shrinkage, weight decay) to control the complexity of the model[22-25].

C. Normalization

Normalization constitutes the final stage of digital preprocessing prior to multivariatepattern analysis. Normalization techniques can be broadly grouped in two classes: local and global methods. Local methods operate across the sensor array on each individual "sniff" in order to compensate for sample-to-sample variations caused by analyte concentration and sensor drift, among others[13,14].

The most widely used local method is *vector normalization* in which the feature vector of each individual sniff is divided by its norm and, as a result, forced to lie on a hypersphereof unit radius is given by

$$y_S^{(k} = \frac{x_S^{(k}}{\sqrt{\sum_S (x_S^{(k)})^2}}$$

$x_S^{(k}$ denotes the response of sensor 's' to the k-th sample in the database. vector normalization can therefore be used to compensate for sample-to-sample variations in concentration. In this context, vector normalization can be applied in situations when each odor has a unique concentration, but discrimination is to be performed on the basis of odor quality. Conversely, this method should not be used when the vector amplitude is known to carry relevant information[12].

Global methods, on the other hand, operate across the entire database for a single sensor (e.g., the complete history of each sensor), and are generally employed to compensate for differences in sensor scaling. Two global procedures are commonly employed in e-nose systems: (i) *sensor autoscaling*, in which the mean and standard deviation of each feature are set to zero and one, respectively,

$$y_s^{(k} = \frac{x_s^{(k} - mean[x_s]}{std[x_s]}$$

and (ii) *sensor normalization*,in which the range of values for each individual feature is setto [0,1] is given by,

$$y_s^{(k} = \frac{x_s^{(k} - \min_{\forall k}[x_s^{(k)}]}{\max_{\forall k}[x_s^{(k)}] - \min_{\forall k}[x_s^{(k)}]}$$

Global methods are typically used to ensure that sensor magnitudes are comparable, preventing subsequent pattern-recognition procedures from being overwhelmed by sensors with arbitrarily large values. For instance, nearest-neighbors procedures are extremely sensitive to feature weighting, and multilayer perceptron's can saturate their sigmoidal activation functions for large inputs [20, 22].

Sensor normalization makes full use of the input dynamic range but, it is very sensitive to outliers since the range is determined by data outliers. Auto scaling, on the other hand, cannot provide tight bounds for the input range but is robust to outliers. However, it must be noted that both techniques can amplify noise since all the sensors (particularly those which may not carry information) are weighted equally [24, 29].

Logarithm metrics have also been used to compensate for highly nonlinear concentration effects. It is also worth mentioning the Box-Cox transform, which could be employed to compensate for nonlinearities, as well as compress the dynamic range of the sensors. It must be noted that these global techniques can amplify noise since all the sensors (including those that may not carry information) are weighted equally. Finally, a *logarithmic* transform can also be used to increase the dynamic range of the system [30,31]

Summary

This paper has presented an overview of the most relevant approaches for signal preprocessing of a sensor response in e-nose system. In baseline manipulation, fractional method having better performance than other. Subsampling and parameter-extraction methods are more commonly employed. Sensorauto scaling can be used to normalize the sensor response and it is robust to outliers.

Conclusion

Although curve fitting method is having lossless compression, but it is computationally intensive. In baseline manipulation, Difference method only eliminates additive errors. Relative method cannot suitable for removing additive errors. Vector normalization method should not be used when the vector amplitude is known to carry relevant information. To conclude, we believe that the development of signal preprocessing techniques for an e-nose system make agrand challenge in future.

References

1. Enobong Bassey, Jacqueline Whalley and Philip Sallis, "An Evaluation of Smoothing Filters for Gas Sensor Signal Cleaning", The Fourth International Comference on Advanced Communications and Computation: INFOCOMP 2014.
2. Miguel Macías, J. Enrique Agudo , Antonio García Manso , Carlos Javier García Orellana, Horacio Manuel González Velasco and Ramón Gallardo Caballero, "A Compact and Low Cost Electronic Nose for Aroma Detection" Sensors 2013.
3. Rehman, A. and Saba, T. (2014). Evaluation of artificial intelligent techniques to secure information in enterprises, Artificial Intelligence Review, vol. 42(4), pp. 1029-1044, doi. 10.1007/s10462-012-9372-9.
4. Saba, T. and Altameem, A. (2013) Analysis of vision based systems to detect real time goal events in soccer videos, Applied Artificial Intelligence, vol. 27(7), pp. 656-667, doi. 10.1080/08839514.2013.787779
5. Rehman, A. and Saba, T. (2012). Off-line cursive script recognition: current advances, comparisons and remaining problems, Artificial Intelligence Review, Vol. 37(4), pp.261-268. doi. 10.1007/s10462-011-9229-7.
6. Haron, H. Rehman, A. Adi, DIS, Lim, S.P. and Saba, T.(2012). Parameterization method on B-Spline curve. mathematical problems in engineering, vol. 2012, doi:10.1155/2012/640472.
7. Saba, T. Rehman, A. Elarbi-Boudihir, M. (2014). Methods And Strategies On Off-Line Cursive Touched Characters Segmentation: A Directional Review, Artificial Intelligence Review vol. 42 (4), pp. 1047-1066. doi 10.1007/s10462-011-9271-5
8. Rehman, A. and Saba, T. (2011). Document skew estimation and correction: analysis of techniques, common problems and possible solutions Applied Artificial Intelligence, Vol. 25(9), pp. 769-787. doi. 10.1080/08839514.2011.607009
9. Saba, T., Rehman, A., and Sulong, G. (2011) Improved statistical features for cursive character recognition" International Journal of Innovative Computing, Information and Control (IJICIC) vol. 7(9), pp. 5211-5224
10. Muhsin; Z.F. Rehman, A.; Altameem, A.; Saba, A.; Uddin, M. (2014). Improved quadtree image segmentation approach to region information. the imaging science journal, vol. 62(1), pp. 56-62, doi. http://dx.doi.org/10.1179/1743131X13Y.0000000063.
11. Neamah, K. Mohamad, D. Saba, T. Rehman, A. (2014). Discriminative features mining for offline handwritten signature verification, 3D Research vol. 5(3), doi. 10.1007/s13319-013-0002-3
12. Saba T, Al-Zahrani S, Rehman A. (2012) Expert system for offline clinical guidelines and treatment Life Sci Journal, vol.9(4), pp. 2639-2658.
13. Mundher, M. Muhamad, D. Rehman, A. Saba, T. Kausar, F. (2014) Digital watermarking for images security using discrete slantlet transform, Applied Mathematics and Information Sciences Vol. 8(6), pp. 2823-2830, doi.10.12785/amis/080618.
14. Norouzi, A. Rahim, MSM, Altameem, A. Saba, T. Rada, A.E. Rehman, A. and Uddin, M. (2014) Medical image segmentation methods, algorithms, and applications IETE Technical Review, Vol.31(3), doi. 10.1080/02564602.2014.906861.
15. Haron,H. Rehman, A. Wulandhari, L.A. Saba, T. (2011). Improved vertex chain code based mapping

algorithm for curve length estimation, Journal of Computer Science, Vol. 7(5), pp. 736-743, doi. 10.3844/jcssp.2011.736.743.

16. Rehman, A. Mohammad, D. Sulong, G. Saba, T.(2009). Simple and effective techniques for core-region detection and slant correction in offline script recognition Proceedings of IEEE International Conference on Signal and Image Processing Applications (ICSIPA'09), pp. 15-20.
17. Selamat, A. Phetchanchai, C. Saba, T. Rehman, A. (2010). Index financial time series based on zigzag-perceptually important points, Journal of Computer Science, vol. 6(12), 1389-1395, doi. 10.3844/jcssp.2010.1389.1395
18. Rehman, A. Kurniawan, F. Saba, T. (2011) An automatic approach for line detection and removal without smash-up characters, The Imaging Science Journal, vol. 59(3), pp. 177-182, doi. 10.1179/136821910X12863758415649
19. Saba, T. Rehman, A. Sulong, G. (2011) Cursive script segmentation with neural confidence, International Journal of Innovative Computing and Information Control (IJICIC), vol. 7(7), pp. 1-10.
20. Rehman, A. Alqahtani, S. Altameem, A. Saba, T. (2014) Virtual machine security challenges: case studies, International Journal of Machine Learning and Cybernetics vol. 5(5), pp. 729-742, doi. 10.1007/s13042-013-0166-4.
21. Meethongjan, K. Dzulkifli, M. Rehman, A., Altameem, A., Saba, T. (2013) An intelligent fused approach for face recognition, Journal of Intelligent Systems, vol. 22 (2), pp. 197-212, doi. 10.1515/jisys-2013-0010,
22. Joudaki, S. Mohamad, D. Saba, T. Rehman, A. Al-Rodhaan, M. Al-Dhelaan, A. (2014) Vision-based sign language classification: a directional Review, IETE Technical Review, Vol.31 (5), 383-391, doi. 10.1080/02564602.2014.961576
23. Saba, T. Rehman, A. Altameem, A. Uddin, M. (2014) Annotated comparisons of proposed preprocessing techniques for script recognition, Neural Computing and Applications Vol. 25(6), pp. 1337-1347 , doi. 10.1007/s00521-014-1618-9
24. JWJ Lung, MSH Salam, A Rehman, MSM Rahim, T Saba (2014) Fuzzy phoneme classification using multi-speaker vocal tract length normalization, IETE Technical Review, vol. 31 (2), pp. 128-136, doi. 10.1080/02564602.2014.892669
25. Saba, T., Rehman, A., Sulong, G. (2010). An intelligent approach to image denoising, Journal of Theoretical and Applied Information Technology, vol. 17 (2), 32-36.
26. MSM Rahim, SAM Isa, A Rehman, T Saba (2013) Evaluation of adaptive subdivision method on mobile device, 3D Research 4 (2), pp. 1-10
27. Ahmad, AM., Sulong, G., Rehman, A., Alkawaz,MH., Saba, T. (2014) Data Hiding Based on Improved Exploiting Modification Direction Method and Huffman Coding, Journal of Intelligent Systems, vol. 23 (4), pp. 451-459, doi. 10.1515/jisys-2014-0007
28. Harouni, M., Rahim,MSM., Al-Rodhaan,M., Saba, T., Rehman, A., Al-Dhelaan, A. (2014) Online Persian/Arabic script classification without contextual information, The Imaging Science Journal, vol. 62(8), pp. 437-448, doi. 10.1179/1743131X14Y.0000000083
29. Rehman, A., and Saba, T. (2013) An intelligent model for visual scene analysis and compression, International Arab Journal of Information Technology, vol.10(13), pp. 126-136
30. Yan, Fengchun Tian, "Feature Extraction From Sensor Data For Detection Wound Pathogen Based Electronic Nose", Sensors and Materials Vol-24,2012.
31. Mazlina Mamat, Salina Abdul Samad and Mahammad A. Hannan, "An Electronic Nose for Reliable Measurement and Correct Classification of Beverages", Sensors 2011.

11/5/2014