Metrol. Meas. Syst., Vol. XXI (2014), No. 4, pp. 649–662.

METROLOGY AND MEASUREMENT SYSTEMS
Index 330930, ISSN 0860-8229
www.metrology.pg.gda.pl

DIFFERENTIAL ELECTRONIC NOSE IN ON-LINE DYNAMIC MEASUREMENTS

S. Osowski[1,2], K. Siwek[1], T. Grzywacz[1], K. Brudzewski[3]

1) Warsaw University of Technology, Faculty of Electrical Engineering, Koszykowa 75, Warsaw, Poland (✉ sto@iem.pw.edu.pl, +48222347235)
2) Military University of Technology, Faculty of Electronic Engineering, Kaliskiego 1, Warsaw, Poland
3) Warsaw University of Technology, Faculty of Chemistry, Noakowskiego 3, Warsaw, Poland

Abstract

The paper presents application of differential electronic nose in the dynamic (on-line) volatile measurement. First we compare the classical nose employing only one sensor array and its extension in the differential form containing two sensor arrays working in differential mode. We show that differential nose performs better at changing environmental conditions, especially the temperature, and well performs in the dynamic mode of operation. We show its application in recognition of different brands of tobacco.

Keywords: electronic nose, differential system, volatile recognition.

1. Introduction

The stability of operation of an electronic nose (e-nose) is an important problem in practical applications, since the temporal drift of operating characteristics of semiconductor sensors due to changes of environmental parameters or internal changes of the semiconductors, causes changes of their signals and, as a result, an improper recognition of volatiles. The important factors influencing the reaction of sensors are the temperature and humidity of the surrounding atmosphere [1–2]. This change can be partly reduced by using an electronic thermostat or by applying a special arrangement of the measurement process.

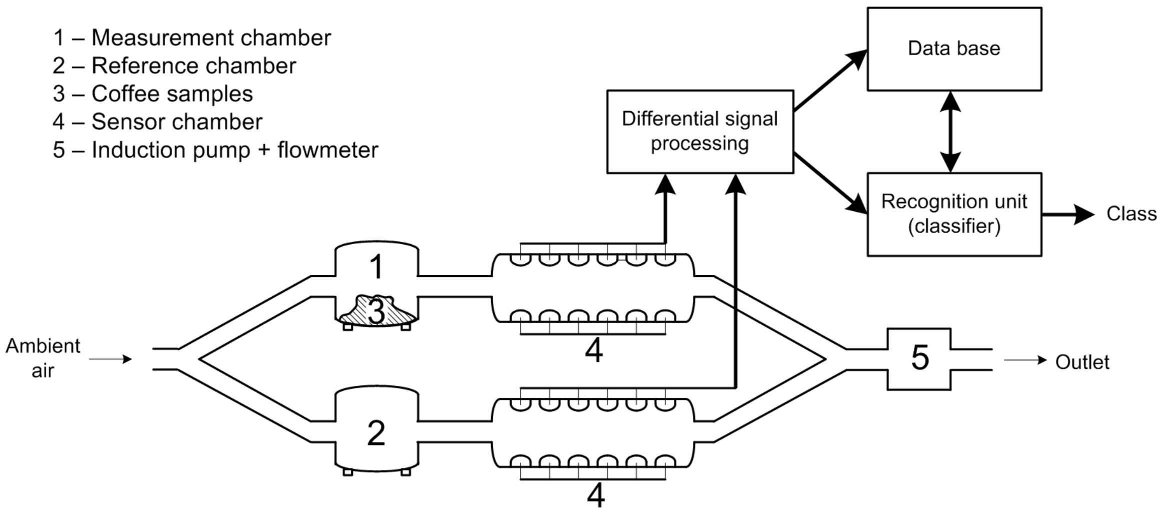


Fig. 1. The general scheme of a differential nose.

Article history: received on Feb. 03, 2014; accepted on Jul. 25, 2014; available online on Dec. 15, 2014; DOI: 10.2478/mms-2014-0053.

The differential e-nose [3−4] of the structure presented in Fig. 1 is an example of such an improvement. It employs two identical channels and identical sensor matrices working in the differential mode. The investigated material is placed in the measurement chamber. The reference chamber is empty. The use of the same ambient air in two channels of e-nose is practical and allows to considerably limit the effect of the instability of the humidity and temperature in the measurement process.

Application of the differential e-nose allows to eliminate the problem of baseline calculation and use it in an on-line mode of operation. At stable environmental measurement conditions the results of the differential nose are equivalent to the classical (one sensor array) e-nose solution. At changing temperature or humidity the differential e-nose allows to make measurements in the on-line mode delivering acceptable results. This is not true in the case of the classical e-nose, since at changing environmental conditions we have to repeat the calculation of the baseline, interrupting the on-line measurement process. Thanks to the differential mode of operation it will be possible to extend the typical application of e-noses [5] to the cases, where a dynamic on-line measurement is needed.

2. Extension of classical e-nose to differential mode

The classical e-nose structure uses one array of few gas semiconductor sensors reacting in a different way to the presence of volatiles. The measured voltage signals of the sensors (proportional to their resistances) are used to generate the features which form the pattern used by a recognizing device (for example a neural network) to perform its classification or regression task. Let us denote the averaged temporal sensor resistance of j-th sensor in the array by $R(j)$. In order to produce the consistent data for the pattern recognition process, we have to eliminate the baseline (the reference value). The baseline values are the measured signals of sensors in the synthetic air atmosphere. As a diagnostic feature we apply the relative variation $r_c(j)$ of each sensor resistance

$$r_c(j) = \frac{R_m(j) - R_0(j)}{R_0(j)}, \tag{1}$$

where $R_m(j)$ is the actual measured resistance of the j-th sensor in the array and $R_0(j)$ represents the baseline value for this sensor. Application of the expression (1) provides automatic normalization of the signals in a classical mode operation of the e-nose. In practice, the value of $R_0(j)$ for the j-th sensor is determined by averaging many measurement points at the synthetic air, which are acquired within the measurement window at determined environmental conditions.

The numerator in the expression (1) represents a difference between the actual resistance of the sensor in the presence of volatiles and the baseline value of this resistance. In this sense the classical (common) mode is close to the differential one. The main drawback of this mode of operation is that the measurements of the analyzed signals and the reference signals (baseline) are collected in different moments of time. In stationary processes this is not a problem. However, in a dynamic mode of operation, when the environmental conditions (temperature, humidity) are changing significantly, the baseline conditions differ from the actual ones. So, the predetermined reference value is not well suited for the actual measurement and leads to errors.

The remedy to such a problem of the e-nose is the simultaneous application of two identical and independent arrays of sensors working in a differential mode. One of them is used as the measurement matrix and the second as the reference one [3]. The measurement array of sensors is exposed to the investigated volatiles and the reference array to the ambient air only. Both arrays are placed close to each other, so the environmental conditions (temperature, humidity, pressure) are approximately the same in both sensor chambers. The

Metrol. Meas. Syst., Vol. XXI (2014), No. 4, pp. 649–662.

streams of sample signals generated by both arrays are then subtracted. In this way the differential signal is produced:

$$R_d(j) = R_m(j) - R_r(j), \tag{2}$$

for j=1, 2, …, N, where index m refers to the measurement and r – to the reference array. Taking into account that both arrays work in the same environmental conditions we observe that the differential signal reduces greatly the common interference effects, such as changes of temperature and humidity of the ambient air. In this system the reference signals correspond to the ambient air conditions used instead of the synthetic air in the classical mode.

Observe that the differential e-nose is automatically calibrated at changing conditions of the ambient air, because the reference and measurement signals are acquired simultaneously in the same atmospheric environments. These features of the differential e-nose allow to use it as an instrument for the applications in dynamically changing conditions.

3. Comparison of classical and differential e-noses

3.1 Theoretical relations

The relations between the classical and differential mode of operation of electronic noses may be compared on the common basis. Let us assume that the measurement array of the differential nose serves in both modes (in the common mode the reference array is simply ignored). The measured signal $R_m(j)$ of j-th sensor in the common mode is described by

$$R_m(j) = R_0(j) + r_c(j)R_0(j), \tag{3}$$

where $R_0(j)$ represents the mean baseline value (the average of many measurements made in the synthetic air) and $r_c(j)$ is the normalized resistance in the common mode. On the other hand, the measured signal of the pair of sensors in the differential mode is described by

$$R_m(j) = R_d(j) + R_r(j). \tag{4}$$

Since the same measurement array is used in both types of measurement we may combine both equations together and get

$$R_d(j) = r_c(j)R_0(j) + \left(R_0(j) - R_r(j)\right). \tag{5}$$

In steady state conditions, the mean value $R_0(j)$ is very close to the instantaneous measurement of the reference resistance $R_r(j)$ in the differential mode. In such a case we can assume

$$R_d(j) \cong r_c(j)R_0(j). \tag{6}$$

It means that in steady state conditions both measurements (in the common and differential modes) are equivalent to each other with respect to the numerical results and the scale of similarity is equal $R_0(j)$. To get the corresponding results in both measurements we have to start from the common mode of operation of the e-nose, calculate the baseline and then switch the system to the differential mode. The advantage of using the differential nose is its operation in the on-line measurement mode (no need for the baseline calculation at changing environmental conditions). It means that the measurements can be done in the dynamic state of the process at changing temperature or humidity.

In the case of instantaneous measurements performed in highly changing environmental conditions (the temperature, pressure or humidity) the results of a classical arrangement of sensors will be inaccurate, since the baseline is calculated in different conditions. Contrary to this, the differential nose is not affected by such a limitation since the reference measurements are performed simultaneously with the acquisition of signals in the measurement chamber.

a. Experimental set-up

To verify this observation we have made experiments of recognition of Robusta and Arabica, the two botanical varieties of coffee. In the numerical experiments we have used the same differential system setup for the differential and common modes. The sensing system was composed of two chemo sensor arrays (the measurement and reference arrays) built of the same types of sensors [3], placed in two independent chambers located nearby. Both sensor arrays are identical and composed of 12 heated metal oxide gas sensors of Figaro series: 2xTGS2600-B00, 2xTGS2602-C00, TGS2610-C00, TGS2610-D00, TGS2611-C00, TGS2611-E00, 2xTGS2612-D00, 2xTGS2620-C00. The sensors of 26xx series were applied, since this family is known from its high stability of operation and small size. Because of a limited number of sensor types, some of them in both arrays have been duplicated at different loadings. The varied loading has been arranged through the potentiometer R wired in series with the sensor, used to tune the level of the sensor output signal in the clean air environment. The sensor resistance R_s in such an arrangement is equal to $R_s = R\frac{V_c}{V_R} - R$, where V_c is the circuit voltage and V_R the voltage of an additional resistor. Four sensors have been duplicated in this way using the additional resistances. Thus, the sensor signals have been changed up to 25%.

Two additional sensors: the temperature sensor LM35DH and humidity sensor HIH-3610-02 have been applied to provide the information on the temperature and humidity. The detailed description of construction of the differential e-nose is given in [3].

Both sensor arrays work in practically the same environmental conditions. Inside the measurement channel there are volatiles of the investigated material coming from its sample amount. The reference channel is supplied only from the outside air and is free from the analyzed volatiles. In the differential mode only the difference of signals is stored and processed further. In the common mode operation of the e-nose only the signals of the measurement array are registered, whereas the reference signals are ignored. Instead, the additional measurements performed in the measurement chamber at the presence of the surrounding air are performed to calculate the baseline.

The measurement array sensed the volatiles of 10g of either Arabica or Robusta placed in a cylindrical vial of the channel. The total volume of this vial was about 200 ml. The second array (the reference one) sensed only the ambient air. In both channels the ambient air was sucked with the flow rate of 0.5 SLPM by an induction pump.

Two types of experiments have been performed. In the first case the temperature as well as the humidity inside the chambers were approximately constant and equal T=28°C and Rh=35% (the steady state ambient conditions). In the second set of experiments we have changed the temperature by heating the air around chambers from 28°C to 36°C. The measurements of the coffee specimen were done 100 times. The registration of the samples in the form of voltages proportional to the sensor resistances was performed every 0.4 seconds. The baseline calculation in the common mode of operation was done only once at the beginning of experiments, at the starting temperature.

Metrol. Meas. Syst., Vol. XXI (2014), No. 4, pp. 649–662.

b. The numerical results

In the first set of experiments we have provided an approximately constant temperature of the air and made experiments in the differential and common modes of operation of the electronic nose. To compare both modes of operation of the e-nose we repeated measurements twice: first in the common mode, including the baseline calculation in the ambient air, and next in the differential mode. All measurements were carried out using the 10g samples of coffee (pure Arabica and pure Robusta). The measurements performed in the ambient air defined the baseline signals of the sensors for the common mode operation at the humidity of 35% and the temperature of 28°C (the laboratory conditions). To get the correct results of comparison of both modes of operation it is important to preserve the order of measurements: first the calibration of the measurement sensor array, then the common mode measurements of the coffee and then the differential mode of measurements of the same aroma.

Fig. 2 presents exemplary numerical results of comparison for the samples of Robusta coffee. Similar results have been obtained for Arabica. There are presented the signals of two chosen sensors in the common mode of operation (indicated by a solid line for sensor 3 and a dashed line for sensor 4) and the differential signals of the same two sensors in the differential mode. Then we have mapped the differential signals to the common mode by applying the relation (6) which is rewritten here in the form

$$r_c^{'}(j) = R_d(j) / R_0(j)\,. \tag{7}$$

The transformed values representing the mapped differential signals are indicated in Figure 2 by the red dots. It is evident that the mapped differential samples match very well the curve of the signals in the common mode operation.

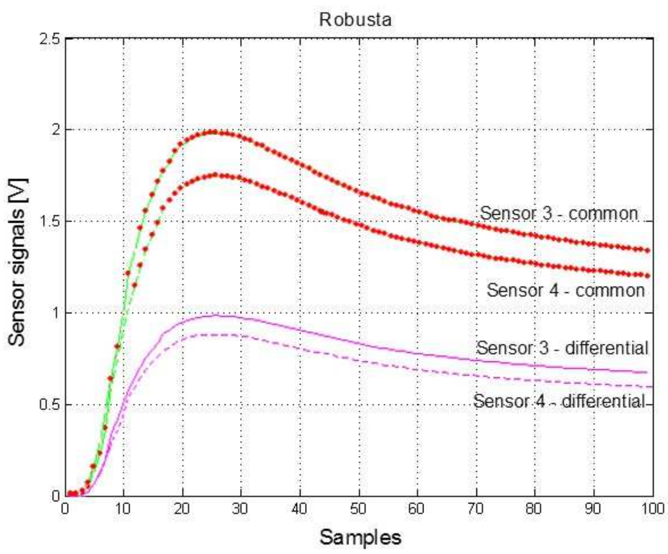


Fig. 2. The results of measurements of Robusta in the common mode (green lines) and differential mode of operations (magenta lines) for two sensor pairs S3=TGS2602 and S4=TGS2600. The red dots represent the mapped values of the differential signals according to the relation (7). The horizontal axis represents the number of subsequent samples (each sample corresponds to 0.4s) and the vertical axis their sensor signals in Volts.

The observed differences between the mapped differential signals and their values in the common mode of operation are negligible, confirming our statement that the common and differential modes of operation of the e-nose are equivalent. Similar results have been obtained in the appropriate measurements of Arabica aroma.

An additional problem is the non-ideality of the same types of sensors forming the pairs. It results in a non-perfect compensation of signals while measuring the same headspace in varying environmental conditions. A typical illustration of this asymmetry is presented in Fig. 3, in which signals of the reference array are presented as negative values according to the equation (2). The signals of measurement and the reference sensor arrays have been registered at the presence of the clean air in both channels. On the basis of many experiments we can state that this effect disturbs the recognition results in a non-significant way.

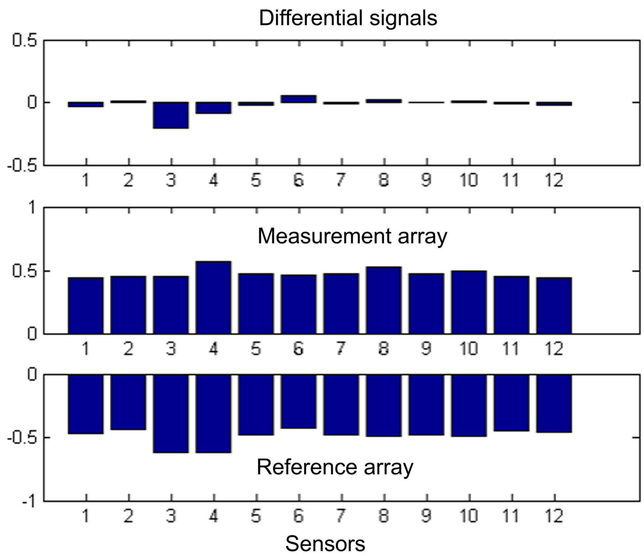


Fig. 3. The illustration of an initial asymmetric effect of the differential e-nose in the clean air measurement. The horizontal axis represents the subsequent sensors and the vertical axis their signals in Volts. The reference signals are presented as negative values according to the equation (2).

The next set of experiments was done in the dynamic mode at quickly changing temperature of an ambient air in the measurement of coffee aroma. We have directly heated the air temperature around both chambers. In these experiments we have kept constant concentration of the coffee (the induction pump was turned off and the airflow in the chambers was blocked). The experiments have been done in the classical and differential modes of operation of the e-nose. The baseline calculation was done only once at the beginning of the experiments at the temperature of 28^{o}C. The humidity level was maintained at a constant level during the whole experiment.

Because of the increasing temperature the sensor signals in the common mode were varying significantly suggesting a change of aroma concentration. In the differential mode of the e-nose operation changes of sensor signals were very small and appropriate to slight differences of temperatures in the measurement and reference channels. The differential mode of operation has shown significant insensitivity of the e-nose to the changing temperature.

Fig. 4 presents the change of temperature in the measurement chamber (the upper diagram) and the measured differences of temperatures in both chambers, observed in the dynamic mode of operation of the e-nose at the presence of Arabica aroma. A slight difference of temperature in both chambers resulted from an imperfect symmetry in heating of both channels in the dynamic mode.

Metrol. Meas. Syst., Vol. XXI (2014), No. 4, pp. 649–662.

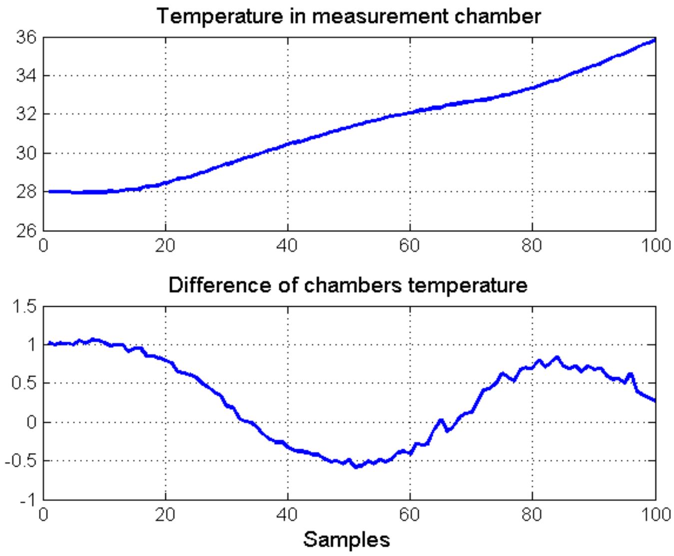


Fig. 4. The change of temperature of the measurement chamber (the upper subplot) and the difference of temperatures in both chambers (the bottom subplot) in the dynamic mode of measurements of the Arabica. The vertical axes are scaled in Celsius degrees.

Fig. 5 shows signals of 4 sensors at changing temperatures (as pointed in Fig. 4) in the common and differential modes of operation of the e-nose in the measurement of Arabica aroma. The upper subplot presents the differential signals and the bottom one - the common mode of operation. Only the differential mode reflects the true (constant) concentration of aroma. In spite of quickly changing temperature the differential signals of the sensor pairs remain on an almost constant level appropriate to the constant concentration of the coffee aroma.

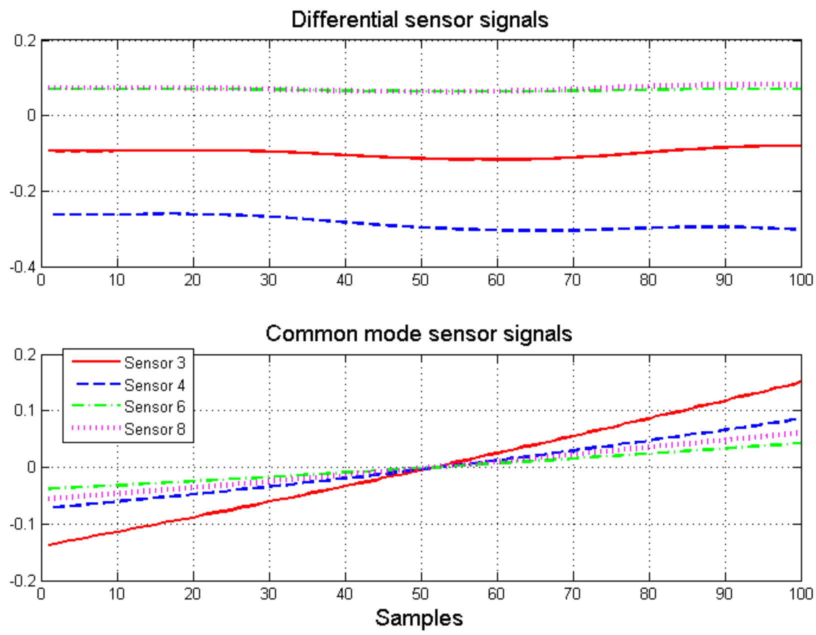


Fig. 5. The change of sensor signals in the differential mode of operation (the upper diagram) and in the common mode of operation (the bottom diagram) in the measurement of Arabica at the changing temperature. The results correspond to the temperature plot shown in Fig. 4. The horizontal axes represent the number of the subsequent samples of sensor signals (each sample corresponds to 0.4s) and the vertical axis the values of these signals in Volts.

Slight changes of the signal values are due to some observed differences of temperature in both chambers, which are inevitable in the dynamic mode of operation. The similar effects were observed also for Robusta. It can be seen that the common mode of operation of the e-nose is useless in the dynamic mode, since it is impossible to verify the baseline at quickly changing temperatures without interrupting the measurements. This experiment proves the advantage of using the differential nose in changing environmental conditions.

4. Application of differential e-nose in recognition of cigarette

4.1 Materials

The developed differential system was applied in recognition of cigarette brands (classes) on the basis of smell of their leaves [6]. The registered and normalized differential signals of the sensor arrays form the input to a support vector machine (SVM) used as the final recognition and classification tool. The leaves of 11 brands of cigarettes, each of the total mass of about 3g, were obtained from the local suppliers. They have been placed in the measurement chamber of the e-nose. The cigarette brands and the country of their production which took part in experiments are depicted in column 1 of Table 1, (EU - the European Union, Russia and Ukraine). Column 2 presents the notation of classes (in abbreviated form) representing different cigarette brands.

Table 1. The cigarette brands and their short notation.

Cigarette brand	**Notation**
Chesterfield Blue (EU)	C1
Chesterfield Red (EU)	C2
Classic Red (Russia)	C3
Fest (Ukraine)	C4
LD Blue (EU)	C5
LD Red (EU)	C6
LD Silver (EU)	C7
Magnat (Ukraine)	C8
Viceroy Blue (EU)	C9
Viceroy Blue (Ukraine)	C10
Viceroy Red (Ukraine)	C11

The investigated brands of cigarettes were produced by different companies in different countries. For example Viceroy Blue was produced in EU and Ukraine and it is interesting to know whether the system is able to recognize them. The basic measurements were done at constant ambient temperature equal to 22^{o}C and humidity 44%. The samples have been registered every 0.5 seconds. The total number of acquired samples was 200 (the total measurement time equal to 100 seconds). Because of the dynamic on-line mode of signal registration the concentration of the gas was changing within the measurement window. The changes of voltage signals of two sensor pairs in the measurement of 11 cigarette brands are shown in Fig. 6.

Metrol. Meas. Syst., Vol. XXI (2014), No. 4, pp. 649–662.

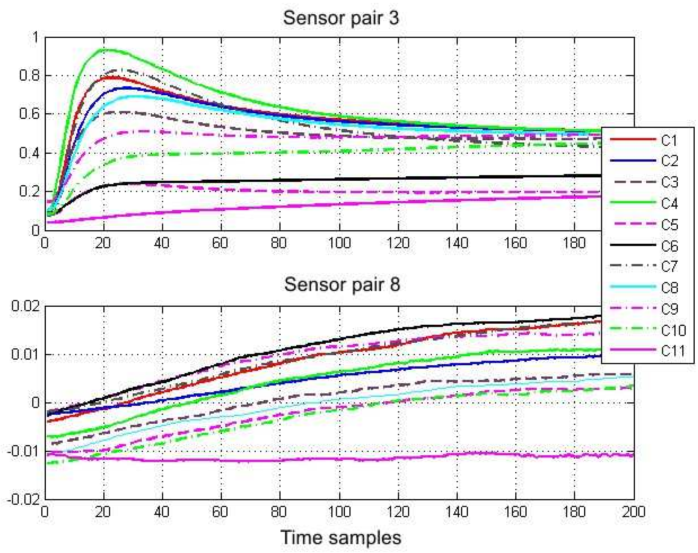
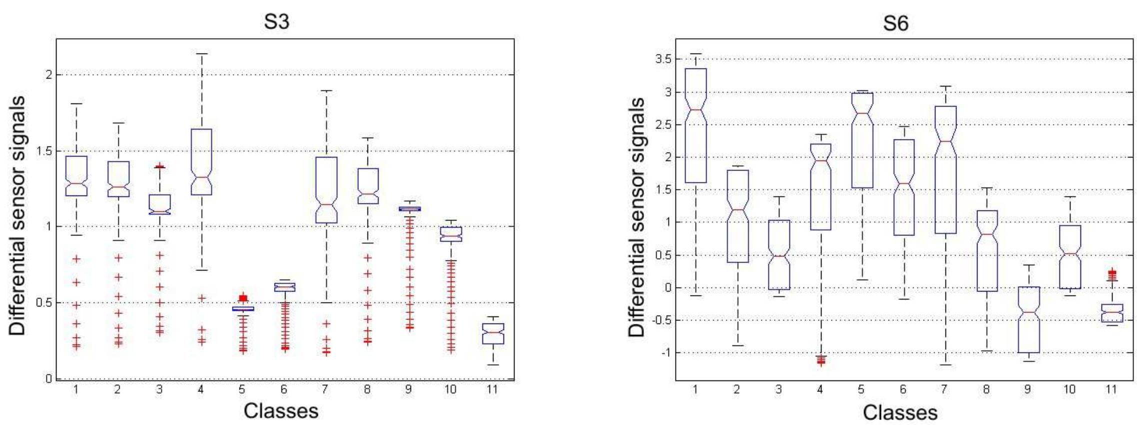


Fig. 6. The change of dynamic responses of two chosen sensor pairs in the measurement of 11 types of smells. The horizontal axes represent the number of subsequent samples of sensor signals and the vertical axis the values of these signals in Volts.

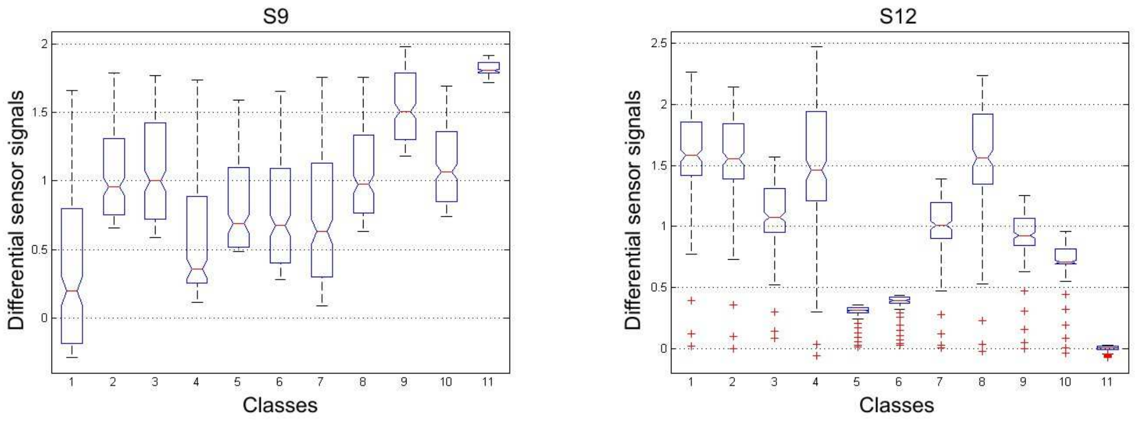
The observed changes of the signals result from the dynamic reactions of sensors to the presence of different tobacco smells. All sensor pairs form the patterns characteristic for the smell of each cigarette brand. Thanks to this each cigarette brand can be associated with the characteristic pattern at any point of time.

4.2 Analysis of sensor signals

An important point in the pattern recognition is to find if the measured values of sensor signals have differing distribution for different classes. We have applied here the test of independence based on the analysis of variance of sensor signals (ANOVA analysis [7]) for all brands of cigarettes. In each case we have got p-value very close to zero (p is the probability that the sensor signals representing different classes come from the same population – the so called null hypothesis). This small p-value means that the null hypothesis should be rejected. Fig. 7 presents the box plots corresponding to the differential sensor signals for 4 chosen pairs of sensors: S3, S6, S9 and S12.



a) b)



c) d)

Fig. 7. The box plots of the sensor signals in the measurement of 11 cigarette brands: a) sensor pair S3, b) sensor pair S6, c) sensor pair S9, d) sensor pair S12. The horizontal axes represent the subsequent classes of cigarettes and the vertical axes indicate the range of change of the differential sensor signals.

Each box in Fig. 7 has lines at the lower quartile, median, and upper quartile values. Whiskers extend from each end of the box to the adjacent values in the data. By default the most extreme values within 1.5 times the interquartile range from the ends of the box. The points displayed with a red + sign denote outliers, that is the data with values beyond the ends of the whiskers. In most cases the signals of the sensors occupy different ranges of values, have different median, and we observe also outliers. If one sensor responses in the same way for the particular classes of volatile compounds (for example sensor S3 for classes 1 and 2) the other (for example S6 or S9) differentiate these classes in a significant way. Such a cooperation of all sensors enables to recognize different classes. The significant difference was observed also for the same type of cigarette (Viceroy Blue) produced in two different countries.

The important information on the separability of classes represented by the sensor signal patterns can be drawn from mapping the 12-dimensional data of the smell pattern into the 2-dimensional space provided by the two most important principal components in a principal component analysis (PCA). The PCA represents a classical statistical technique for analyzing the covariance of the multivariate statistical observations related to sensor signals [6],[8]. It reveals the structure behind the correlation of many variables and is described as the linear transformation $\mathbf{y}=\mathbf{W}\mathbf{x}$ mapping the N-dimensional original feature vector $\mathbf{x}$ into the K-dimensional output vector $\mathbf{y}$ of $K<N$. The vector $\mathbf{y}$ preserves the most important elements of original information and $\mathbf{W}$ is the PCA transformation matrix composed of the eigenvectors of the correlation matrix $\mathbf{R}_{xx}$ associated with the set of feature vectors $\mathbf{x}_i$. Mapping the 12-dimensional vectors into 2-dimensional space allows to trace the trajectory of changing values of many sensor signals within the measurement process of all cigarette brands. Fig. 8 presents the 2-dimensional PCA plot of 200 samples representing each cigarette class. The horizontal axis represents the first principal component PC_1 and the vertical the second one - PC_2. The first component carries 95.2% of total information and the second 3.26%.

Metrol. Meas. Syst., Vol. XXI (2014), No. 4, pp. 649–662.

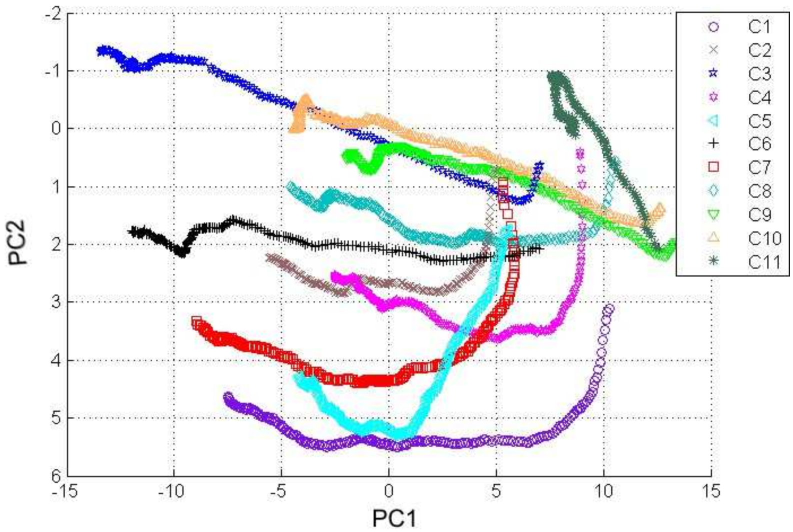


Fig. 8. The dynamic trajectories of two most important principal components of the differential sensor signal measurement of all 11 cigarette smells.

It is evident that the data corresponding to different classes of cigarettes are grouped in the separate clusters forming the shape of a continuous trajectory reflecting the dynamic mode of measurements. The presented results show reasonably good separation of data in most of the space regions. However, there are some restricted areas of space where the data of different classes are mixing.

4.3 Classification

The 12-dimensional sensor signal pattern is used in the second step to recognize smells and associate them with the proper brand of cigarettes. As the classifier we use here the support vector machine (SVM) network of Gaussian kernel having the reputation of the most efficient classification tool [8].

The SVM is a one-output linear machine working in the high dimensional feature space formed by nonlinear mapping of the original N-dimensional input vector $\mathbf{x}$ into a K-dimensional feature space ($K>N$) through the use of the kernel functions $K(\mathbf{x},\mathbf{x}_i)$. We will use here the Gaussian function $K(\mathbf{x},\mathbf{x}_i)=\exp\left(-\gamma\|\mathbf{x}-\mathbf{x}_i\|^2\right)$. The most important advantage of the SVM over other neural solutions is the fact that its learning algorithm is based on the quadratic programming with linear constraints. The primary learning problem of the SVM is formulated as the task of separating training vectors $\mathbf{x}_i$ into two classes of the destination values, either d_i=1 or d_i=-1, with the maximum separation margin between classes. The large width provides immunity of the system to the existence of noise and other artifacts in the testing samples.

The constant value of the regularization parameter C, responsible for the minimization of errors on the learning data was applied. It is an important parameter, because it controls the tradeoff between the complexity of the machine and the number of non-separable data points used in learning. A small value of C results in acceptation of more not separated learning points. For a higher value of C we get a lower number of classification errors on the learning data points, but a more complex network structure and potentially worse generalization ability. The optimal values of C and parameter γ are determined after additional series of learning experiments through the use of the validation data set (10% of learning data). In the learning process many different values of C and γ were used. The optimal values are the ones for which the classification error on the validation data set is the smallest. In this way we have got γ=1 and C=1000.

In the case of many classes we use a strategy called "one against one" [8]. In this approach at M classes we train $M(M-1)/2$ individual SVM networks to recognize among all combinations of two classes of data. Then the vector **x** belongs to the class of the highest number of winnings in all combinations of 2-class SVM networks. For M=11 classes we have to train 55 SVM sub-networks.

The input vector **x** in our classification problem was composed of the differential signals of the sensor arrays. Their values were normalized by dividing each column through the maximum absolute entry of this column. In the case of 12 sensors applied in the measurement the size of input vector **x** is also 12. The destination vector **d** associated with the input vector represents the class to which the appropriate data belongs. We recognized 11 classes of cigarettes, as is shown in Table 1.

The available data used in the experiments were composed of 200 patterns corresponding to each class (altogether 2200 data pairs). To obtain the most objective assessment of the developed classification system we have applied the cross-validation approach. The randomly selected data were split into 2 equal parts. One of them was used in learning and the second in testing. The procedure of learning and testing was repeated 50 times and each time the learning and testing parts were randomly chosen.

The testing error on the data not used in learning was calculated as their average in all 50 runs on the testing data. Its observed mean value was 0.30% at the standard deviation equal to 0.28%. This is a very good result, especially taking into account that all measurements have been obtained in the dynamic mode, without applying a lengthy baseline determination procedure.

Table 2. The confusion matrix at recognition of 11 classes of cigarettes in the 10-fold cross-validation mode.

Class	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
C1	99.8	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C2	0.0	99.3	0.1	0.0	0.0	0.0	0.4	0.0	0.2	0.0	0.0
C3	0.1	0.0	99.8	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
C4	0.0	0.0	0.2	99.2	0.0	0.0	0.0	0.6	0.0	0.0	0.0
C5	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
C6	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0
C7	0.0	0.5	0.0	0.0	0.0	0.0	99.5	0.0	0.0	0.0	0.0
C8	0.0	0.0	0.1	0.2	0.0	0.0	0.0	99.5	0.0	0.0	0.2
C9	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	99.8	0.0	0.0
C10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.9	0.1
C11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	99.9

Table 2 presents the confusion matrix obtained in 50 cross-validation experiments. It is represented in the relative terms. The diagonal entries represent the percentage of properly recognized classes of cigarettes. Each entry outside the diagonal means the relative error. The entry in the (i,j)-th position of the matrix means the false assignment of i-th class to the j-th one.

From the analysis of this matrix it is evident that the nonzero values of the outside diagonal terms are very scarce. There is only a limited number of misclassifications indicated by the nonzero elements outside the diagonal. Classes C5 and C6 have been recognized perfectly in all runs. The highest percentage of misclassifications have been observed for class C4, but even in this case the relative mean error was equal only 0.8%. The confusion matrix confirms that practically all classes have been recognized with the acceptable accuracy. An interesting fact is that the most often confused classes are C4 and C8. Both cigarette brands

Metrol. Meas. Syst., Vol. XXI (2014), No. 4, pp. 649–662.

are produced in Ukraine and it is quite probable that their producers used similar tobacco sources.

The next experiments with the learned system have been conducted on the data registered in the new runs of the system at slightly different conditions (the ambient temperature changing from 20^{o}C to 24^{o}C and humidity changing from 30% to 40%). These experiments have done on different days. Each day new samples of the measured materials have been supplied. We have noted slightly different sensor signals (the average relative difference in 200 samples was below 15%). The system trained on the original data set was tested on the newly acquired data. The average results of 5 different measurements in the form of the confusion matrix are presented in Table 3. The total relative recognition error was equal to 4.09%.

Table 3. The results of testing the system on the newly acquired data.

Class	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
C1	98.6	0.0	0.6	0.0	0.0	0.1	0.8	0.0	0.0	0.0	0.0
C2	0.1	93.4	0.5	1.0	0.1	0.4	1.3	2.7	0.5	0.1	0.0
C3	0.0	0.0	96.6	0.0	0.0	0.1	0.0	1.3	0.9	0.9	0.2
C4	0.4	0.8	0.5	94.8	0.0	0.1	0.1	2.6	0.1	0.3	0.4
C5	0.1	0.0	0.1	0.0	97.0	2.7	0.1	0.0	0.0	0.0	0.0
C6	0.1	0.0	0.1	0.0	3.8	95.8	0.1	0.0	0.0	0.0	0.0
C7	0.7	0.5	0.2	0.0	0.1	0.7	97.7	0.0	0.1	0.0	0.0
C8	0.1	2.3	1.0	1.6	0.0	0.1	0.1	93.3	0.4	0.4	0.6
C9	0.2	0.0	0.8	0.0	0.0	0.0	0.1	0.1	95.6	3.1	0.0
C10	0.1	0.0	0.8	0.0	0.0	0.0	0.0	0.1	2.8	95.2	1.0
C11	0.3	0.0	0.7	0.0	0.0	0.1	0.0	0.5	0.1	1.2	97.1

This time the highest misclassifications happened in recognition between C9 and C10 classes of cigarettes. They represent the same brand Viceroy Blue produced in two different countries. The lowest accuracy was observed in recognition of the classes C2 (93.4%) and C8 (93.3%). The misclassifications for these two classes were distributed among all other classes.

Another point that should be taken into account is the noise, which is inevitable in the automatic acquisition of data. Its source lays in independent changes of parameters of the sensors resulting from changing the environmental conditions. We have performed the additional numerical experiments by corrupting the measured data using the random noise. Similarly to the results of [9] we have found our SVM recognition system relatively noise-resistant.

5. Conclusions

The analysis of the performance of the differential electronic nose in the dynamic mode of measurements has been studied in the paper. Application of sensors working in the differential mode increases the measurement system sensitivity and makes it less susceptible to the change of the environmental conditions, because these changes in the differential signal are to some degree compensated. Moreover, we avoid a lengthy procedure of the baseline determination which makes the system suitable for on-line measurement applications.

The other benefits of using it are essential in the dynamically changing environmental conditions. The classical mode of operation needs baseline estimation in the actual conditions, which makes it impossible in on-line measurements. The differential nose has no such restrictions. We have proved its advantage over the classical arrangement of sensors at the temperature changing quickly in the process of measurement.

The differential nose has proved its advantages in recognition of 11 brands of cigarettes. The measurements of tobacco volatiles made on-line in the dynamic mode have shown that the differential e-nose is capable to recognize the cigarette smells very quickly and with a high accuracy.

References

[1] Korotcenkov, G., Cho, B. K. (2011). Instability of metal oxide-based gas sensors and approaches to stability improvement (short survey). *Sensors and Actuators B*, 156, 527−538.

[2] Gardner, J. W., Bartlett, P.N. (1999). *Electronic noses – principles and applications*. Oxford: Oxford University Press.

[3] Brudzewski, K., Osowski, S., Ulaczyk, J. (2010). Differential electronic nose of two chemo sensor arrays for odor discrimination. *Sensors and Actuators B*, 145, 246−249.

[4] Kwon, H. J., Kim, D. G., Hong, K. S. (2013). Multiple odor recognition and source direction estimation with an electronic nose system. *International Journal of Distributed Sensor Networks,* ID 361378, 1−7.

[5] Wilson, D., Baietto, M. (2009). Applications and advances in electronic-nose technologies. *Sensors,* 9, 5099−5148.

[6] Brudzewski, K., Osowski, S., Golembiecka, A. (2012). Differential electronic nose and Support Vector Machine for fast recognition of tobacco. *Expert Systems with Applications*, 39, 9886−9891.

[7] Matlab Signal Processing Toolbox (2013). Natick: Mathworks.

[8] Schölkopf, B., Smola, A. (2002). Learning with kernels. Cambridge, MA: MIT Press.

[9] Barman, I., Dingari, N. C., Singh, G. P., Soares, J. S., Dasari, R. R., Smulko, J. M. (2012). Investigation of noise-induced instabilities in quantitative biological spectroscopy and its implications for noninvasive glucose monitoring. *Analytical Chemistry*, 84(19), 8149−8156.