Methodological questions of intellectual multisensor analyzers construction

Zhitkov A.N.1, Fedyukhin A.V.2, Filaretov G.F.1, Chervova A.A.1, O.V. Derevianko3

1National Research University «Moscow Power Engineering Institute», 111250 Russia, Moscow, Krasnokazarmennaya 14.

2V. A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Russia, Moscow, Profsoyuznaya, 65.

3Peter the Great St. Petersburg Polytechnic University, 194021, Russia, St.Petersburg, Polytechnicheskaya 29

gefefi@yandex.ru, fedyukhinav@yandex.ru

Abstract. In this paper we consider actual questions, connected with construction and usage of multisensor analyzers of "electronic nose" or "electronic tongue" types. Despite a large amount of papers concerning various aspects of this theme, there is a significant lack of methodological works among them. This work presents a discussion on general key problems of improvement of intellectual multisensor analyzers, namely the problems of optimization of sensor system, selection of algorithm means, evaluation of identification ability of the analyzer, typification of the utilized project design solutions at its construction. Typical examples of intellectual multisensor analyzers are presented, which are successfully used in various areas, including that for control and monitoring of water quality.

Introduction

Intellectual multisensor analyzers are referred to scientific devices of the new generation, which are based on combined usage both of achievements in microelectronics, chemical material science, sensors, and modern information technologies. The main purpose of this type of analyzers is express-analysis of content of complex air or water mediums, express classification (diagnostics) of the analyzed medium state or state of the object under control aiming at determination of its consumer properties [1–11]. Actually, these analyzers form a separate class of information-measuring systems. The main feature of these systems is usage of teaching procedures. This circumstance allows one to say that these systems might be considered as technical means, which simulate olfactive or taste systems of living organisms. That's why they are called "Artificial nose" in case of analysis of air (gas) medium or "Artificial (electronic) tongue" in case of liquid mediums [8].

Despite a large amount of papers concerning various aspects of this theme, there is a significant lack of methodological works among them, which consider general key problems of improvement of intellectual multisensor analyzers. We refer here to the problems of optimization of sensor system, selection of algorithm means, evaluation of identification ability of the analyzer, typification of the utilized project design solutions at its construction. This consideration will be performed using the
results of long-term investigations of the National Research University “MPEI” in the area of methodological and program and algorithmic provision of multisensor systems and its usage primarily for liquid medium analysis [12].

**Synthesis of optimum sensor system**

It should be noted immediately that consumer characteristics of quality identifying devices for various application areas are primarily governed by characteristics of the sensor system (sensor block). Generally, the sensor block contains a massive of sensors of rather high dimension \( n \) (typical \( n \) values are in the range from 10 to 30).

The process of sensor block creation, as a rule, is iteration and includes some stages:

a) Initial synthesis of sensor system. It is carried out on the basis of general considerations of the character of the analyzed medium, its possible chemical components, ions, etc. The obtained from the initial synthesis stage sensor system must be of substantial excess character, i.e. contain a large number of sensors, which sometimes exceed the run-time version.

b) Experimental investigation of the synthesized system using the typical samples of the production which will be further analyzed. The aim of the investigation is to determine operating capability of the system, its principal suitability for solving the considered applied problem. In the case of positive answer to this question, the obtained experimental material might be used at the further stages of sensor block development.

c) Determination of sensor informativeness and choice of its optimum (minimum) amount. It is valuable to carry out this stage using the specially designed technique, which provides two-stage procedure for selection of the most informative sensors [13]. Basing on the results of the first stage, the sensor ranking for its informativeness takes place. The low-sensitive and likely-responsive sensors are drop out. The first stage is devoted to informativeness check of the sensors selected at the first stage by direct usage of classification algorithms, which are laid down in the analyzer.

d) Tests of the optimized sensor set in conditions which are close to practical implementation. At positive result of the tests, one might say that the functional sensor block of the certain analyzer was created.

**Algorithmic provision of the multisensor analyzers**

It is known, that the key feature of the considered multisensor analyzers is that they operate using the principle of preliminary teaching on samples. The end result of the teaching experiment is the training sampling of the volume \( N \). It contains data set came from \( n \) sensors of the sensor block in every experiment with indication of identity of the sample to a certain product (class). In an \( n \)-dimension factor space (sensor space) the elements of the training sampling are displayed as a certain cloud on \( N \) points.

A large number of various classification algorithms are known [14–17]. They are used for identification of separating borders between certain classes basing on the training sampling.

Nowadays in practice the most popular are the following algorithms:

a) The main components method (MCM) in two-dimensional case, when all experimental points are displayed on the plane. Usually elliptical areas are displayed simultaneously, which characterize the degree of its variation for each class (Fig.1). These ellipsis define some attraction areas of points, which are attributed to different classes. They denote, where with probability \( p \) may appear the next point, if it is attributed to this class (as a rule, \( p=0.9 \) or \( p=0.95 \) is chosen).

This method has a good visibility and interpretability of the obtained classification results.

b) Multi-layer perceptron with sigmoidal activation functions and the amount of hidden layers from 1 to 3 and output layer of \( m \) neurons according to the amount of identifying classes. The results are represented as column diagrams, which show the values of output network signals \( y_j, j = 1, 2, \ldots, m \). At this the sample is referred to \( f \) class if the following condition is fulfilled: \( 0.75 \leq y_f \leq 1.0 \), the others \( y_j \leq 0.25 \). If these conditions are not fulfilled, the classification is considered as unreliable.
It should be noted that both these methods have an essential weakness: if at analyzer usage there is a possibility for appearance of new class samples, which are not included into the training sampling, then the whole recognition algorithm should be taught again. Hence the interest for ANN, based on ideas of adaptive resonance theory (ART), which are free of the above-said limitation.

**Evaluation of the identification ability of the multisensor analyzers**

This evaluation is carried out for the purpose of reliability assessment of the identification results, performed by the analyzer during solving a certain applied task. As an identification reliability factor one usually use probability $P$ of right classification of samples of all classes. $P$ value is set by the user, usually $P = 0.95$. Solution of this task might be traced to a standard problem of examination of statistical hypothesis about the value of probability $P_0$ of wrong classification at the given level of criterion $\alpha = 1 - P$. Without getting into details of elementary mathematical transformations, we will write the resulting formula for critical boundaries of amount of cases for wrong classification $N_{\text{fault}}$ to the total amount of experiments $N$, which were carried out during the tests:

$$\frac{N_{\text{fault}}}{N} = u_{1-\alpha} \sqrt{\frac{P_0 (1-P_0)}{N}} + P_0,$$

where $u_{1-\alpha}$ is quantile of normal distribution. This formula is an approximate one, but gives rather precise results at $N>20$. If, for example, $N=25$, $P_0 = P = 0.95$, $u_{0.05} = 1.64$, when using the above formula we will obtain the critical boundary for the amount of faults at identification $N_{\text{fault}}=3.07$. Consequently, if during the tests we revealed more than 3 faults, we conclude that it has insufficient identification ability.

It should be noted that for some applied tasks it is important to check identification ability of the analyzer for revealing of new production types, which were not represented during its teaching.

**Typification of the project design solutions at construction of multisensor analyzers**

Fig.2 presents an example of general view of multisensor analyzer of "Electronic tongue" type, intended for identification of various bottled drinking water types and discovery of counterfeit products (the amount of brands in the initial variant is $m = 6$).
The analyzer consists of a sensor block of 10 ion-selective electrodes, which were selected according to the method described above. Algorithmic provision includes statistical algorithm (MCM with further usage of linear discriminate analysis), neural network algorithm (four-layer perceptron with amount of neurons at the layers $n_1 = n$, $n_2 \approx n/2$, $n_3 \approx n_2 + m/2$, where $n = 10$) and self-teaching ART-algorithm. At combined usage of all three algorithms the decision on attribution to one class or another is made according to the principles of majority logic. The analyzer has successfully passed all necessary tests and further showed high efficiency at solving of tasks for identification of various water mediums.

Besides this, it might be useful for organizing of monitoring process for drinking water quality, which is obtained after filtration of water of poor quality. For this purpose we may recommend to use three samples as initial ones used at the teaching stage: low-quality water from its source; the same water after purification using a certain filter (at initial stage the filtering element is new); and a reference water sample (for example, distilled water). The latter one is used for creation of a reference set on the plane of the first major components for comparison with other samples. The further work of the analyzer in the operation mode allows one to estimate changes in water quality and discover how good is the filtering system operation, as well as how quickly do the filter characteristics change and when the filtering element should be changed or regenerated. It is clearly seen from gradual shift of the representation points which correspond to the filtered water at the plane of two first major components from the initial state (new filtering element) sidewards the points corresponding to water from source, as it is shown in Fig.3.
The idea of typification was also used at construction of other multisensor analyzers. Among them is analyzer of "electronic nose" type for identification of perfume productions, multisensor analyzer of vodka authenticity (MAP-V), which received the Certificate of Russia Gosstandard for confirmation of measurement type with registration in State Register of Measuring Equipment.

Conclusions

For successful implementation of intellectual multisensor analyzers one should clearly understand its place and role among the other techniques and analytical equipment intended for solving of these tasks. They can't replace the organoleptic tests, chromatographic analysis, mass-spectrometry methods, etc. Its major advantage is relatively low cost of analysis, usage simplicity, objectivity, rapidness and absence of sample preparation. So up to now its main purpose is express-analysis, preliminary rejection of production according to its quality, including field tests. It is clear that in the course of gaining experience and increasing of its possibilities, the area of its application will gradually increase.

References

[1] Yaroshenko I, Kirsanov D, Kartsova L, Sidorova A, Borisova I and Legin A 2015 Determination of urine ionic composition with potentiometric multisensor system Talanta 131 556–61
[2] Debus B, Kirsanov D, Yaroshenko I, Sidorova A, Piven A and Legin A 2015 Two low-cost digital camera-based platforms for quantitative creatinine analysis in urine Anal. Chim. Acta 895 71–9
[3] Yaroshenko I, Kirsanov D, Kartsova L, Sidorova A, Sun Q, Wan H, He Y, Wang P and Legin A 2016 Exploring bitterness of traditional Chinese medicine samples by potentiometric electronic tongue and by capillary electrophoresis and liquid chromatography coupled to UV detection Talanta 152 105–11
[4] Shkodyrev V P and Yagaf arov K I 2018 The approach to emergency situation prediction in dynamical systems using neural networks vol 2018–Febru (Association for Computing Machinery) pp 27–32
[5] Averin I A, Igoshina S E, Karmanov A A, Pronin I A, Moshnikov V A and Terukov E I 2017 Simulation of the sensor response of vacuummeters with sensitive elements based on multicomponent oxide nanomaterials with the fractal structure Tech. Phys. 62 799–806
[6] Baryshnikova M V, Filatov L A, Petrov A S and Alexandrov S E 2015 CVD Deposited Titania Thin Films for Gas Sensors with Improved Operating Characteristics Chem. Vap. Depos. 21 327–33
[7] Mileshko L P, Kotov V N, Korolev A N, Cherepahin I I and Scherbinin I P 2009 Multisensor systems: textbook (Taganrog: Izd-vo TTI YUFU)
[8] Vlasov Y G, Legin A V and Rudnickaya A M 2008 Electronic tongue is a system of chemical sensors for water analysis Ros. him. zh. LIII 101–12
[9] Filaretov G F and Zhitkov A N 1999 Application of artificial neural networks in sensor systems Datchiki i Sist. 5 2–9
[10] Kashaev R S, Kozelkov O V. and Kubango B E 2017 On-line flow proton magnetic resonance analyzers for oil-well liquid control on GOST P 8.615–2005 GSI Proc. High. Educ. institutions. ENERGY Sect. Probl. 19 137–51
[11] Kashaev R S, Kozelkov O V., Malev N A and Kubango B E 2017 Automatic control of rotary pump using measurements of oil-well liquid by on-line proton magnetic resonance analyzer Proc. High. Educ. institutions. ENERGY Sect. Probl. 19 119–31
[12] Filaretov G F and Zhitkov A N 2001 Analyzer of “Artificial nose” type Materialy XIII nauchno-tekhnicheskoy konferencii “Datchiki i preobrazovateli informacii sistem izmereniya, kontrolya i upravleniya” (Datichi – 2001) (MGIEHM) pp 301–3
[13] Filaretov G F 2001 Method for synthesis of sensor block of analyzer of “Artificial nose” type Materialy XIII nauchno-tekhnicheskoy konferencii “Datchiki i preobrazovateli informacii
sistem izmereniya, kontrolya i upravleniya” (Datchik – 2001) (MGIEHM) pp 303–5

[14] Utkin L V, Chekh A I and Zhuk Y A 2016 Binary classification SVM-based algorithms with interval-valued training data using triangular and Epanechnikov kernels Neural Networks 80 53–66

[15] Sonkin K, Stankevich L, Khomenko Y, Nagornova Z, Shemyakina N, Koval A and Perets D 2016 Neurological classifier committee based on artificial neural networks and support vector machine for single-trial EEG signal decoding 9719 100–7

[16] Stankevich L and Sonkin K 2016 Human-robot interaction using brain-computer interface based on EEG signal decoding 9812 LNCS 99–106

[17] Utkin L V and Zhuk Y A 2017 Interval SVM-Based Classification Algorithm Using the Uncertainty Trick Int. J. Artif. Intell. Tools 26