PARETO-OPTIMUM APPROACH TO MATHEMATICAL MODELING OF ODOURS IDENTIFICATION SYSTEM

Andriy Zavorotnyy, Veda Kasyanyuk

***Abstract**: Mathematical model of vapor identification system is developed. Calibrating signals from vapor sensors are used to identify unknown input to vapor sensors and approximate output from eventual sensor system. Approximation formulas are resulted from pareto-optimum solution of multi-criterion problem. The developed method can be used to create new measuring-calculating systems within "device + PC = device with added benefits" framework.*

***Keywords**: identification, an odorant, impacted data, measuring system, pareto-optimization*

***ACM Classification Keywords**: I.6 Simulation and Modeling*

Introduction

Modeling of awareness is one of the most important tasks scientists have faced in the space of human modeling. This problem is complex and comprises lots of subtasks that can be considered as pretty much independent problems though. The most outstanding problems are investigation of decision-making process, modeling of human memory, pattern identification, etc. Furthermore the last one is deemed as a complex problem and consists of identification of visual images («an electronic eye»), identification of liquids («electronic tongue»), recognition of odours («an electronic nose»), etc. In this article one of possible approaches to modeling of "an electronic nose» is presented.

Measurements Methodology

Structure of gas media can be placed on record by means of quartz microbalance. This approach is founded on the fact of proportion between weight of substance adsorbed on a surface of a quartz plate (lets denote it as Δm) and frequency of fluctuations of the quartz (lets denote it as Δf) [Eichelbaum, 1999]:

$$\Delta f \approx -\frac{2f_0^2}{A\sqrt{\rho_q \mu_q}} \Delta m \ ,$$

Here f_0 is a fundamental frequency of the quartz resonator, A is an effective area of a surface of quartz, μ_q is a “piezoelectric hardness” and ρ_q is a density of the quartz. Selectivity of sensor system is achieved due to coating of quartz surface with layers that are sensitive to given chemical elements.

Measurements process with sensor-based system results into output signals data set. This data then can be interpreted and unknown odor can be identified within reasonable timeframe. For that mathematical method has been developed. This method take into account that output signals can be inaccurate. It is also assumed that sensor system is complex enough therefore its model is deemed to be unknown. Instead the method uses sensor

system's measurements of template odors to assess unknown odor and simulate output of eventual measuring system for this odor.

Mathematical Model of Measurements Results Processing

The proposed method of mathematical processing of sensors' signals is based on pareto-optimum approach to calibration of unknown model with etalon measurements results [Belov, 2001].

According to the approach the unknown model of measuring process can be presented in terms of linear operator that is defined on a Hilbert space. A series of template measurements should be conducted. Then the results of measurements are used to approximate just unknown input to measuring process, i.e. without interim approximation of unknown model of the process. In this way ultimate approximation results are less dependent on interim errors that probably were present should unknown model would be approximated.

Approximation is posed as an optimization problem with two criteria. The first criterion is minimization of noise energy, i.e. dispersion of noise in optimization result. The second criterion is minimization of population mean of squared subtraction of expected output signal from approximated output signal across all template inputs.

The optimization problem is set according to pareto-optimum principle, i.e. these two criteria are directed towards minimum simultaneously.

So, considering the primary assumption that communication between an input and an exit from sensors system is described by the unknown linear operator and considering presence of noise in results of measurements, we can write down

The following mathematical model describes measuring process of input odor by sensor-based system

$$\vec{y} = G\vec{u} + \vec{\nu}\,. \tag{1}$$

The model (1) takes into account that measuring process is linear since G is unknown linear operator that is mathematical abstraction of sensor system. Furthermore the model consider known measurement results $\vec{y} = (y_1 \dots y_n)^*$ are impacted by environment noise $\vec{\nu} = (\nu_1 \dots \nu_n)^*$. Unknown input $\vec{u} = (u_1 \dots u_m)^*$ is mathematical abstraction of a multi-component odor that should be identified. It's a set of content levels of primary odorants. Asterisk * denotes an adjunction of an element, i.e. either vector or matrix.

Let's assume that $M(\vec{\nu}) = 0$, where M denotes an average of distribution. Then let's $\Re$ denotes an operator of covariance of noise [Pitiev, 1989]. The operator $\Re$ is deemed to be known and nonsingular. Finally let's consider q template measurements that have been conducted according to the scheme (1) on known odorants $\vec{u}_j$, $j = \overline{1,q}$. As a result we have q measurement results

$$\vec{y}_j = G\vec{u}_j + \vec{\nu}_j,\ j = \overline{1,q}\,. \tag{1}$$

For all measurements (2) nose summands $\vec{\nu}_j$, $j = \overline{1,q}$ are deemed to satisfy to $M(\vec{\nu}_j) = 0$ condition and result into the same covariance operator $\Re$.

Once it's assumed that sensor system's parameters are unknown let's use measurement results (2) to resolve optimization problem that has been set with criteria described above. Should P denotes eventual sensor-based

system that provide desirable processing of the same unknown odor then the optimization objective is to approximate both unknown content levels of primary odorants $\hat{\vec{u}}$ and $\hat{P\vec{u}}$. P is deemed to be set beforehand, e.g. by odor identification domain expert.

Approximation of $P\vec{u}$ is done by means of processing with a linear operator B of known signal $\vec{y} = (y_1 \dots y_n)^*$ [Pitiev, 1989]. The operator B should be calculated in the way to satisfy pareto-optimization problem stated above. In formal terms the problem is set as follows

$$\begin{cases} h(B) = M\|B\vec{\nu}\|^2 \to \min\limits_B \\ \varphi(B) = M\sum\limits_{j=1}^{q}\|B\vec{y}_j - P\vec{u}_j\|^2 \to \min\limits_B \end{cases} \quad (3)$$

The first criterion in (3) poses minimization of noise energy. Respectively the second criterion is a minimization of mean of squared subtraction of expected output signal from approximated output signal.

This problem has been put and successfully resolved in the general case for linear operators G, P, B that are defined on a Hilbert space [Zavorotnyy, 2004]. Additional restriction of being bounded has been imposed to the operator G.

To resolve Pareto problem (3) let us minimize convex convolution

$$M\left(\lambda\|B\vec{\nu}\|^2 + (1-\lambda)\sum_{j=1}^{q}\|B\vec{y}_j - P\vec{u}_j\|^2\right) \to \min_B,\ \lambda \in (0;1). \quad (4)$$

To find operator B that would be optimum in reference to (4) Frechet derivative from convex convolution (4) should be set equal to zero. The constructed equation can be solved and result into continuum set of solutions in the form

$$B(\alpha) = \sum_{j=1}^{q} P\vec{u}_j f_j^*\left(\sum_{j=1}^{q} f_j f_j^* + (\alpha + q)\Re\right)^{-1},\ \alpha \in (0;1). \quad (5)$$

In formula (5) conventional signs mean $f_j = M\vec{y}_j$ and $\alpha = \dfrac{\lambda}{1-\lambda}$ is a pareto-optimisation parameter. Any operator B in (5) is an effective solution of Pareto problem (3).

The developed method has been used for interpretation of measurements results to identify odorants from spirit group. Therefore the equation (2) and the problem (3) have been instantiated in terms of Euclidean spaces. Specifically operator $\Re$ has been defined as covariant matrix R. Similarly to the assumption that has been stated for $\Re$ the matrix R has been deemed non-singular too, i.e. $\det R \neq 0$. In turn operator P has been instantiated as identity operator. Therefore content levels of primary odorants have been approximated.

It is easy to see from (5) that in this case formulas for input and output of sensor-based system (i.e. $\hat{\vec{u}}$ and $\hat{\vec{y}}$) are in the form:

$$\hat{\vec{u}} = \sum_{j=1}^{q} \vec{u}_j f_j^* \left(\sum_{j=1}^{q} f_j f_j^* + (\alpha + q)\Re \right)^{-1} \vec{y}, \ \alpha \in (0;1), \tag{6}$$

$$\hat{\vec{y}} = \aleph \hat{\vec{u}},$$

In the expression (6) $\aleph$ denotes an operator that is destined to model desirable processing of an odor. Usually such operator is defined by domain knowledge holder.

For approximate calculation of $\hat{f}_j = M\vec{y}_j$ in (6) statistical average can be used. It's quite common practical approach.

Pareto-optimization parameter α should be used for parity regulation between criteria of pareto-optimisation problem (3). This two criteria have opposite trends, i.e. while α increases the first criterion is reducing whereas the second criterion is increasing to $\sum_{j=1}^{q} \left\| Pu_j \right\|^2$. It's shown in [Zavorotnyy, 2004] that the criteria stick to conservation low $\varphi'(B(\alpha)) + \alpha h'(B(\alpha)) = 0$.

Indeed in terms of pareto-optimum problem (3) every single solution (5) has no benefits or disadvantages setting it against to any other solution (5). However in practice specific value of parameter α should be used. Let's present some common approaches to select α based on general principls of multi-criteria optimization [5].

For instance α can be found from minimization of sum of criteria, i.e. $\alpha = \arg\min_{\alpha} (\varphi(B(\alpha)) + h(B(\alpha)))$. From derivative of function $z(\alpha) = \varphi(B(\alpha)) + h(B(\alpha))$ it's easy to get stationary point. Taking into account conservation low presented above and the fact that $h'(B(\alpha)) < 0$ [Zavorotnyy, 2004] it's easy to see that $z'(\alpha) = h'(B(\alpha))(1-\alpha) = 0$ and the only possible stationary point is $\alpha = 1$. This is minimal value of $z(\alpha)$ since $z''(\alpha) = h''(B(\alpha))(1-\alpha) - h'(B(\alpha))$ and therefore $z''(1) = -h'(1) > 0$.

"Eldorado" principle can be used to select α too. According to the principle pareto-optimum parameter is resulted from minimization of sum of squared parameters of optimization, i.e. $\alpha = \arg\min_{\alpha} (\varphi^2(B(\alpha)) + h^2(B(\alpha)))$. In other words criteria α is selected as close as possible to coordinate origin in the space formed with all possible values of criteria $\varphi(B(\alpha))$ and $h(B(\alpha))$.

Conclusion

Let's highlight that identification of spirits has been chosen for instance. In the same way it is possible to qualify other odorants either mixes of odorants, also to identify whether an odor belongs to given group. Of course for each case appropriate odor samples are required while calibration measurements.

Therefore the developed approach once applied to output of sensor-based system brings identification to qualitatively new level. It is resistant to data errors. As a result it can be used in real identification systems and increase accuracy of identification. Moreover duy to low amount of computation this approach can be a valued addition to existent sensor-based systems with low impact to their response time. The proposed mathematical

processing can be implemented with and ran by PCs. Therefore it allows saving time and money since there is no need to create specific device with high identification accuracy.

To sum up the quality of volatile compounds identification can increased significantly due to the developed approach of mathematical processing of signals fed from sensor system given it has got optimum set of sensitive coating layers.

Acknowledgements

The paper is published with financial support by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine (www.aduis.com.ua).

Bibliography

[Eichelbaum, 1999] F. Eichelbaum, R. Borngraber, J. Schroder, R. Lucklum and P. Hauptmann, Interface circuits for QCM sensors, Review of Scientific Instruments, 70, may, 1999.

[Belov, 2001] Belov Yu. A., Zavorotnyy A. L., Kasyanyuk V. S. "On one approach to problem of processing of measurements using calibration signals based on multicriteria optimization", Journal of Automation and Information Sciences, 2001, #5 pp. 257-267.

[Pitiev, 1989] Pitiev Yu.P Mathematical methods of experiment interpretation, High school, Moscow, 1989.

[Zavorotnyy, 2004] Zavorotnyy, A. Decision of super-resolution measuring-calculating system modelling problem based on multicriterion optimization. – Bulletin of the University of Kiev, Series: Physics &Mathematics, 2004, #3, pp. 198-205.

Authors' Information

Andriy Zavorotnyy *– PhD, Researcher; TC science sector, theoretical cybernetics department, faculty of cybernetics, Kyiv National Taras Shevchenko University, Glushkov av. 2, build. 6, Kyiv-03127, Ukraine; e-mail: zalbxod@mail.ru*

Major Fields of Scientific Research: pareto-optimization, operator model of measuring-calculating system, fuzzy theory

Veda Kasyanyuk *– PhD, Head of Science Sector, TC science sector, theoretical cybernetics department, faculty of cybernetics, Kyiv National Taras Shevchenko University, Glushkov av. 2, build. 6, Kyiv-03127, Ukraine; e-mail: veda.sia@mail.ru*

Major Fields of Scientific Research: measurements' reduction to calculations, calibration of unknown measurements' model, spline approximation