Analysis of Gas Sensor Response
Using Subspace-Based Prony's Method

Yoshinori Takei,[*] Yasoshima Hiroyuki,[1] Kazuki Hirasawa,
Hidehito Nanto, and Kiyoshi Wada

AMS R&D Center, Kanazawa Institute of Technology, 3-1 Yatsukaho, Hakusan, Ishikawa 924-0838, Japan
[1]Graduate School of Engineering, Kanazawa Institute of Technology,
7-1 Ohgigaoka, Nonoichi, Ishikawa 921-8501, Japan

This paper shows the feature extraction for gas discrimination from the transient response of a gas sensor using a subspace-based system identification technique. We have proposed an approximation and analysis method for the transient response using a subspace-based identification method. This is based on Prony's method and can extract time constants and gain parameters of multicomponents of a sensor output response. The estimated parameters can be used to improve gas discrimination in multivariate analysis as elements of pattern vectors. Transient responses of semiconductor-type gas sensors were evaluated using the proposed method. We show that the method can extract features connected with characteristics of measured gases from these sensor responses.

1. Introduction

Electronic nose (e-nose) systems[1] have been proposed to discriminate gases and odors. The basic form of the e-nose systems can be constructed from a gas sensor array and a discrimination method, *e.g.*, multivariate analysis or machine learning algorithms. The pattern vectors used in the discrimination method, which include information on odors obtained by gas sensors, are prepared. Then amplitude values of the gas sensor array, *e.g.*, steady-state values of the sensors, are ordinarily used as the elements of the vectors. The transient response of the sensors, however, includes more information on odors,[2] and the extracted parameters, *e.g.*, the time constant, can be expected to improve the discrimination of odors in the analysis. Therefore, we have proposed a feature extraction method[3] using a system identification technique, in which Prony's method[4] can be interpreted in the framework of subspace identification, called the multivariable output error state space model identification (MOESP) method.[5,6] We have discussed the efficiency of the method only in numerical simulations.

In this study, we considered a feature extraction of real gas sensor response data in terms of the capability of the proposed method. A semiconductor-type gas sensor was used to evaluate the method. We show that the proposed method can extract feature parameters from the step-wise sensor response data, and we also show a practical procedure to apply the proposed method to actual data.

*Corresponding author: e-mail: takei@neptune.kanazawa-it.ac.jp

2.   Review of Subspace-Based Prony's Method

In this section, we review the subspace-based Prony's method[3] based on the estimation method of the sum of exponential functions as a model of the gas sensor response. Assuming that the interaction between the gases and the surface of the gas sensor can be regarded as a first-order reaction,[2] the transient gas sensor response to the odorants as mixtures can be modeled as the sum of the step response of a discrete first-order lag system. The estimate of the model parameters can be obtained as the parameters of the measurement system. Therefore, the actual gas sensor response can be observed as a high order lag system even if a single component gas is exposed to the sensor, because the measured gas sensor response depends on the pathway of the gas to the sensor, *e.g.*, it includes the bulk diffusion of the gases.

The sum of exponential model can be represented as

$$\tilde{y}_k = \sum_{i=1}^{n} C_i(1 - r_i^k),\tag{1}$$

where

$$r_i = \exp(-\Delta t/\tau_i).\tag{2}$$

The Δt is the sampling period, τ is a time constant, and C is a steady-state value that means the initial concentration or mass adsorbed at saturation. The n in the above equation is the order of the true sensor output model. The dominant parameters that should be extracted, however, cannot be obtained from a model with low order. Therefore, a high order of the model, that is, the model with the order n set large enough, is used to improve the accuracy of the estimates.

We first consider the exponential function fitting in Prony's method[4] to approximate the sensor response by estimating the parameters τ and C in Eq. (1). To obtain τ and C, the sensor output model (1) can be represented as the sum of exponentials. Then, using the following L-step ahead differential,

$$y_k = \tilde{y}_{k+L} - \tilde{y}_k,\tag{3}$$

the sum of exponentials can be introduced to rewrite the output

$$y_k = \sum_{i=1}^{n} m_i r_i^k.\tag{4}$$

For example, the i-th component output $y_{k,i}$ can be shown as

$$\begin{aligned}
y_{k,i} &= \tilde{y}_{k+L,i} - \tilde{y}_{k,i}, \\
&= C_i[1 - \exp(-L\Delta t/\tau_i)] \cdot r_i^k.
\end{aligned}\tag{5}$$

Then L should be chosen to be larger than the largest time constant τ_i to maintain gain sufficiently and can be decided by considering data length obtained in practice.

The represented gain m shown in Eq. (4) and C are related as

$$m_i = C_i[1 - \exp(-L\Delta t/\tau_i)], \quad i = 1, ..., n. \tag{6}$$

Consider the estimations of τ and C through the identification of the model shown in Eq. (4). Equation (4) added the noise term e_k and can be represented as the discrete time state space model

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k, \tag{7}$$

$$y_k = \mathbf{c}^\mathrm{T}\mathbf{x}_k + e_k, \tag{8}$$

where the system matrices $\mathbf{A}$ and $\mathbf{c}$, and the initial state vector $\mathbf{x}_0$, are

$$\mathbf{A} = \begin{bmatrix} r_1 & & & 0 \\ & r_2 & & \\ & & \ddots & \\ 0 & & & r_n \end{bmatrix}, \tag{9}$$

$$\mathbf{c}^\mathrm{T} = [m_1\ m_2\ \cdots\ m_\mathrm{n}], \tag{10}$$

$$\mathbf{x}_0 = [1\ 1\ \cdots\ 1], \tag{11}$$

and the stacked vector defined as follows is introduced:

$$\mathbf{y}_i(k + N) \triangleq [y_{k+N}^\mathrm{T}\ \ y_{k+N+1}^\mathrm{T}\ \ \cdots\ \ y_{k+N+i-1}^\mathrm{T}]^\mathrm{T}, \tag{12}$$

where $\mathbf{y}_i(k + N) \in \mathbb{R}^i$. The term i is the auxiliary order, and N denotes the measured data length. The stacked vector of the noise $\mathbf{e}_i(k + N)$ can be defined similarly. Then the Hankel matrix by using the stacked vectors can be defined as

$$\mathbf{Y}_{k,i,N} \triangleq [\mathbf{y}_i(k)\ \ \mathbf{y}_i(k + 1)\ \ \cdots\ \ \mathbf{y}_i(k + N - 1)]. \tag{13}$$

Corresponding definitions for the noise $\mathbf{E}_{k,i,N}$ are also defined in similar way. In relation to the order of the system n, the pair i and N satisfy $i > n$ and $N \gg n$. Defining the state vector sequence as

$$\mathbf{X}_{k,N} \triangleq [\mathbf{x}_k\ \ \mathbf{x}_{k+1}\ \ \cdots\ \ \mathbf{x}_{k+N-1}], \tag{14}$$

we obtain the following data equation:

$$\mathbf{Y}_{k,i,N} = \mathbf{\Gamma}_i \mathbf{X}_{k,N} + \mathbf{E}_{k,i,N}. \tag{15}$$

The $\mathbf{\Gamma}_i$ called the extended observability matrix is defined as

$$\mathbf{\Gamma}_i = \begin{bmatrix} \mathbf{c}^\mathrm{T} \\ \mathbf{c}^\mathrm{T}\mathbf{A} \\ \vdots \\ \mathbf{c}^\mathrm{T}\mathbf{A}^{i-1} \end{bmatrix}. \tag{16}$$

The key problem of the subspace method is to estimate the column space of $\boldsymbol{\Gamma}_i$ including the system matrices.

To extract the column space of $\boldsymbol{\Gamma}_i$, we compute the singular value decomposition (SVD)[7] of the matrix $\boldsymbol{Y}_{k,i,N}$ given in Eq. (17), i.e.,

$$\boldsymbol{Y}_{k,i,N} = \begin{bmatrix} \boldsymbol{U}_n & \boldsymbol{U}_n^{\perp} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_n & O \\ O & \boldsymbol{\Sigma}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_n^{\mathrm{T}} \\ (\boldsymbol{V}_n^{\perp})^{\mathrm{T}} \end{bmatrix}, \tag{17}$$

where "$\perp$" denotes perpendicular and the matrix $\boldsymbol{\Sigma}_i$ is diagonal. The matrix $\boldsymbol{U}_n \in \mathbb{R}^{i \times n}$ gives the estimate of $\boldsymbol{\Gamma}_i$. The order of the matrix $\boldsymbol{\Sigma}_n$ can be equal to that of the true system. Note that to evaluate the obtained singular values, the order of the system, that is, the number of components of gases can be determined. Using the matrix $\boldsymbol{U}_n$ given in Eq. (17), we solve the following set of equations to estimate the system matrices $\hat{A}$ and $\hat{c}$.

$$\hat{c} = \boldsymbol{U}_n(1,:) \tag{18}$$

$$\boldsymbol{U}_n^{(1)}\hat{A} = \boldsymbol{U}_n^{(2)} \tag{19}$$

In Eq. (18), the submatrix of $\boldsymbol{U}_n$ is denoted using a MATLAB-like notation. $\boldsymbol{U}_n^{(1)}$ is the submatrix composed of the first (i–1) rows of the matrix $\boldsymbol{U}_n$, and $\boldsymbol{U}_n^{(2)}$ is constructed using the last rows in a similar way. Equation (19) expresses the *shift-invariance* property in the estimate $\boldsymbol{U}_n$. The time constant τ can be estimated by applying the eigenvalue decomposition (EVD) to $\hat{A}$.

Once the time constants can be obtained, we can easily estimate the left gain parameter m using the least squares method with the measured data of the sensor response.

3.   Experimental Setup

A commercially available semiconductor-type gas sensor TGS2620 (Figaro Eng. Inc., Japan) was used in experiments because of its low selectivity in response to various gases. Ethanol and other gases were prepared at concentration of 100 ppm. Here we show the results of the response to ethanol only. The gas sensor located in the closed chamber with turbulator was exposed to the gases injected with a microsyringe. Then the sensor output was collected by a notebook PC via an A/D converter.

4.   Results and Discussion

Applying the proposed method to the gas sensor response data, we obtained the estimates to model the gas sensor response under the following condition: user defined parameters were set to L = 800, i = 100, and n = 20. Data length was N = 4200, and the sampling period was 10 ms.

Figure 1 shows the gas sensor response (blue line) measured in the closed chamber with step-wise gas injection and the approximated response (red line) by the proposed method. The result shows that the sensor response can be well reconstructed and that the method can capture features of the sensors response.

$$
\hat{\boldsymbol{\tau}} =
\begin{bmatrix}
7.4102871 \\
0.0026178 \\
0.0004404 + i * 0.0255280 \\
0.0004404 - i * 0.0255280 \\
0.0003377 + i * 0.0106424 \\
0.0003377 - i * 0.0106424 \\
0.0002127 + i * 0.0121852 \\
0.0002127 - i * 0.0121852 \\
0.0000236 + i * 0.0058639 \\
0.0000236 - i * 0.0058639 \\
\vdots
\end{bmatrix}
\tag{20}
$$

$$
\hat{\boldsymbol{C}} =
\begin{bmatrix}
0.2079128 \\
-0.0809885 + i * 1.349 * 10^{-12} \\
-0.0001320 + i * 0.0000818 \\
-0.0001320 - i * 0.0000818 \\
0.0004187 + i * 0.0003294 \\
0.0004187 - i * 0.0003294 \\
0.0000927 + i * 0.0002263 \\
0.0000927 - i * 0.0002263 \\
0.0002727 + i * 0.0002832 \\
0.0002727 - i * 0.0002832 \\
\vdots
\end{bmatrix}
\tag{21}
$$

Estimates of model parameters as shown above are the first ten elements sorted in descending order in terms of the real part of the time constant τ, the parameters of which were obtained up to the defined order of the model n. We can see that the estimate of the time constant τ of the dominant
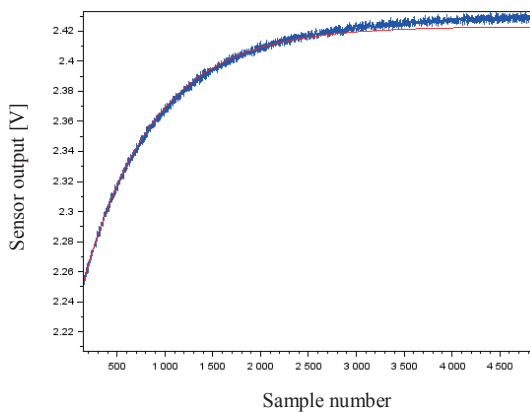

Fig. 1.   (Color online) Transient response of semiconductor-type gas sensor to ethanol gas. The blue line is measured data of the gas sensor response and red line is the approximated response.
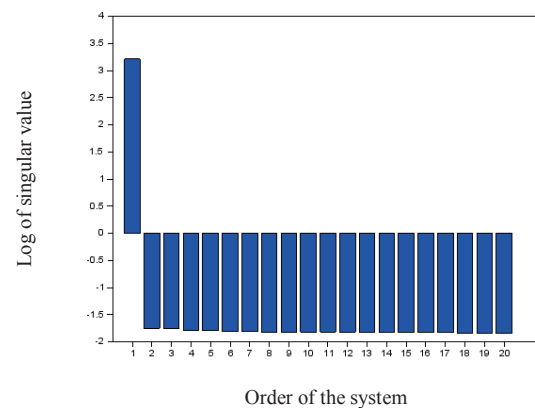
Fig. 2.   (Color online) Results of singular value decomposition to estimate the order of the gas sensor response model.

component can be obtained as 7.41 s. Correspondingly, the saturation value C can be approximated as 0.21 V, which is the change in output value from the initial value. The proposed algorithm includes a SVD as one of numerical tools in the MOESP, which can provide information on the order of the system. In Fig. 2, the singular values are shown, and the number of the relatively large values means the order of the system, that is, the number of gas components. Figure 2 also shows that the gas sensor response can be interpreted as a single gas component.

Therefore, the proposed method can extract the features of the gas sensor response and the extracted features can be used to improve discrimination of gases. Then the estimation procedure using the proposed method can be described as follows.

Extraction of sensor response parameters τ and C
Step 1   Set user-defined parameters L, i, and n.
Step 2   Execute the subspace-based Prony's method and obtain the model parameter pairs $\hat{\tau}$ and $\hat{C}$.
Step 3   Sort the obtained parameter pairs $\hat{\tau}$ with $\hat{C}$ in descending order of the real part of $\hat{\tau}$, and then extract the dominant $\hat{\tau}$ with $\hat{C}$ up to the number of the order decided by using the results of the SVD of the data matrix.

5.   Conclusions

In this paper, we have considered a feature extraction of real gas sensor response data in terms of the capability of the subspace-based Prony's method. Then we have shown that the proposed method can extract feature parameters from the step-wise gas sensor response data and can approximate the transient sensor response. The practical procedures to apply the proposed method to actual data have also be shown, and the method can be used to improve the discrimination of gases.

References

1   T. C. Pearce, S. S. Schiffman, H. T. Nagle, and J. W. Gardner: Hand-book of Machine Olfaction (Wiley-VCH, Weinheim, 2003).
2   S. Seto, H. Kawabe, L. Shi, Y. Shimomura, T. Oyabu, and T. Katsube: Sens. Mater. **18** (2006) 1.
3   Y. Takei, H. Nanto, and K. Wada: Sens. Mater. **26** (2014) 163.
4   Y. Takei, S. Koyama, H. Nanto, and T. Mihara: Sens. Mater. **21** (2009) 209.
5   M. Verhaegen and P. Dewilde: Int. J. Control **56** (1992) 1187.
6   M. Verhaegen: Int. J. Control **58** (1993) 555.
7   G. H. Golub and C. F. Van Loan: Matrix Computations, 3rd ed. (Johns Hopkins University Press, Baltimore and London, 1991).