Japanese Journal of Applied Physics

PROGRESS REVIEW

Odor sensing system with multi-dimensional data analysis

To cite this article: Takamichi Nakamoto 2019 *Jpn. J. Appl. Phys.* **58** SB0804

Japanese Journal of Applied Physics **58**, SB0804 (2019) **PROGRESS REVIEW**
https://doi.org/**10.7567/1347-4065/ab0740**

Odor sensing system with multi-dimensional data analysis

Takamichi Nakamoto*

Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama, 226-8503, Japan

*E-mail: nakamoto@nt.pi.titech.ac.jp

Received November 1, 2018; accepted February 7, 2019; published online April 1, 2019

There are two current approaches in odor sensing systems. One is an odor biosensor using actual olfactory receptors. Although there have been many artificial odor sensors proposed over the past three decades, there is still room for improvements in terms of sensitivity, selectivity and stability. A biomimetic approach using olfactory receptors of a living body is expected to enhance capability. The second is a deep learning technique to predict odor impressions. The mapping of mass spectra onto sensory test data based on the semantic differential method was performed. A method of using two autoencoders for both independent- and dependent-variable spaces together with ordinal multi-layer perceptrons was proposed. Its classification accuracy was better than that of the conventional regression method. Moreover, several cost functions were applied to the autoencoder and evaluated. It was found that the Itakura–Saito cost function was superior to others for reproducing small peaks in the mass spectrum. The author believes these are key technologies for realizing a sophisticated olfactory sensor.

1. Introduction

In our olfactory system, an output pattern of various types of olfactory receptor neurons with partially overlapping specificities is recognized by an olfactory neuron system.[1)] Thus, a biomimetic approach to realizing an artificial sensing system using an array of sensors followed by pattern recognition has been proposed.[2,3)] This type of odor sensing system is called an electronic nose.[4)] Many applications such as food, beverage, cosmetics, health care, and environmental testing are expected if a sensing system with sufficient performance is available.

A variety of sensors, such as metal oxide gas sensors,[5)] MOSFET gas sensors,[6)] conducting polymer sensors,[7)] surface acoustic wave sensors,[8)] cantilever sensors,[9)] quartz crystal microbalance sensors,[10)] optical gas sensors,[11)] electrochemical gas sensors[12)] etc., have been studied for three decades. However, sensitivity, selectivity and stability need to be further improved. Since a living body has excellent olfactory capability, a biologically-inspired approach is very promising.[13–15)] High capability of olfaction is expected when an OR (Olfactory Receptor) can be used as a sensor element. Thus, we introduce a biomimetic odor sensing system using cells expressing plural types of olfactory receptors.[16)]

Next, a deep learning technique[17)] to predict odor impression was studied.[18)] Although many researchers have studied odor classification and sometimes odor quantification, there are just a few works related to the prediction of odor impression. The prediction of odor impression is a relatively new area. Rich expression like in humans is expected if a machine can say whether a sample smells sweet, fresh, fruity etc. This is important when we realize human olfactory interfaces.

In our system, mass spectrum data were mapped onto the results of sensory testing using a large-scale neural network. The mass spectrum is used here since its stability and reproducibility are high enough to form the mapping function. Moreover, a neural network to predict a wide range of data was studied since small peaks in a mass spectrum have much to contribute to predictions.[19)] Recent results relating to the odor sensing system are explained in this review paper. This paper is an extension of the abstract from Solid State Devices and Materials 2018,[20)] which describes the methods of this new odor sensing system.

2. Odor biosensor

The principle of odor recognition in the living body is shown in Fig. 1. Three ORs respond to specific stereo-chemical structures such as the rectangle, triangle and circle as represented in this example. Although each OR responds to multiple odorants, its output pattern from an array of ORs is unique. Thus, odorants are recognized using a pattern recognition technique. Although this technique is also used for artificial sensing systems, the selectivity is still not sufficient even if a combination of sensor array and pattern recognition is used. Thus, we study odor biosensors.

We use insect ORs because of their simple structures and mechanism. Sf21 cells expressing Drosophila ORs were used here. Since fluorescent protein such as GCaMP6s inside the cell is sensitive to calcium ions, the increase in calcium ion concentration inside the cell caused by the ion channel opening due to odorant reception changes the intensity of fluorescent light.[21)] In the case of GCaMP6s, the fluorescent light is green (488 nm), whereas excitation light from a laser diode is blue (510 nm). Figures 2(a) and 2(b) show the fluorescent image of an Sf21 cell expressing OR56a before and after geosmin injection. Geosmin has a typical moldy smell and its detection is important for drinking water and food. The fluorescent intensity increased after geosmin injection as OR56a captures geosmin. Thus, the change of fluorescent light intensity is used as a sensor output.

Fluorescent images of the cells were taken by CMOS camera though a dichroic mirror as shown in Fig. 3. A sample flow system was used to switch the sample solution and linger solution using a three-way solenoid valve, repeatedly. Either the sample or linger solution was supplied to the cell in a sensor chamber. An image sensor (OV-7740, Omni Vision) was used to capture fluorescent images. Since the fluorescent light is weak, it takes some time to accumulate the light signal. Thus, the frame rate of the image was 10 frames per second. The image was transferred to Field Programmable Gate Array (5ASTFD5K3F40I3, Altera).

In our measurement system, a lock-in technique was used to detect fluorescent light synchronous with the modulated

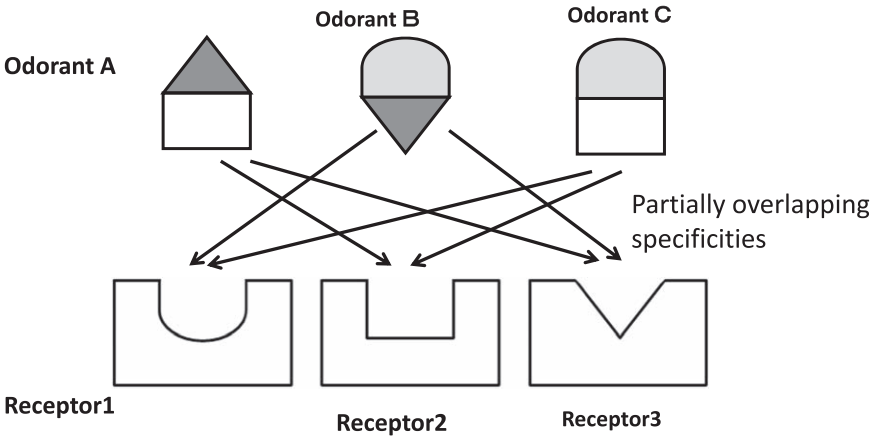

Fig. 1. Principle of odor recognition.

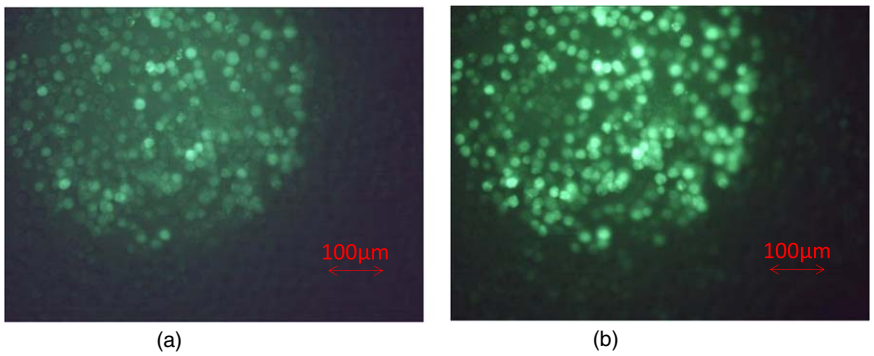

Fig. 2. (Color online) Fluorescent images of Sf21 cell expressing OR56a before geosmin injection (a) and after its injection (b).

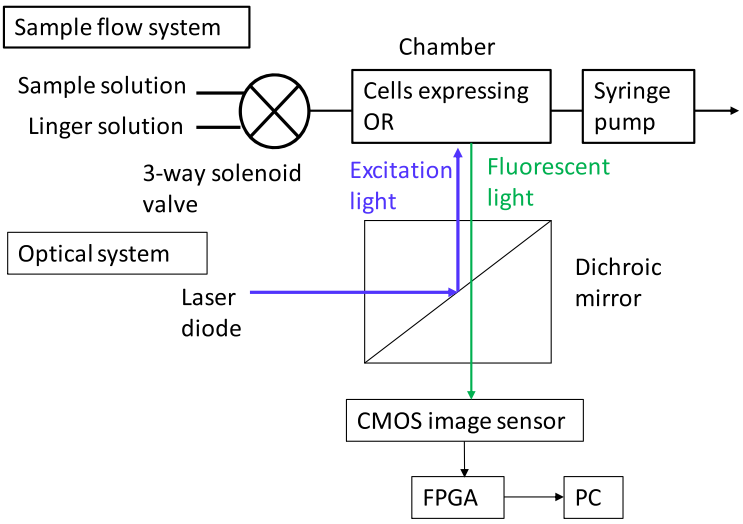

Fig. 3. (Color online) Experimental setup for the odor biosensor.

light from a laser diode.[22,23] A schematic of lock-in measurement is shown in Fig. 4. The ambient light asynchronous with the modulated excitation light can be removed. Moreover, the signal to noise ratio can be raised using this technique. Since the fluorescent light is very weak, the lock-in technique is useful. The digital lock-in amplifier was implemented into FPGA as mentioned above. It consists of a BPF (Band Pass Filter), PSD (Phase Sensitive Detector) and LPF (Low Pass Filter). Since the modulation frequency was 1 Hz, the center frequency of the BPF was also 1 Hz. The cutoff frequency of the LPF was 0.1 Hz since the current response time was relatively slow at about 20–30 s.

These circuits were designed using Matlab/Simulink (Mathwork) and were then converted into Hardware Description Language.[24] The lock-in amplifiers work in parallel. Although one lock-in amplifier per pixel is ideal, the amount of circuit required is too large. Thus, 100 lock-in amplifiers were implemented into FPGA. The image was divided into 100 areas and the intensity in each area was integrated so that it could be input to the lock-in amplifier.

After the experiment on a single OR, the cells with two types of OR were randomly distributed in a chamber. Although the cell patterning technique was not required in our method, the image recognition technique was used to

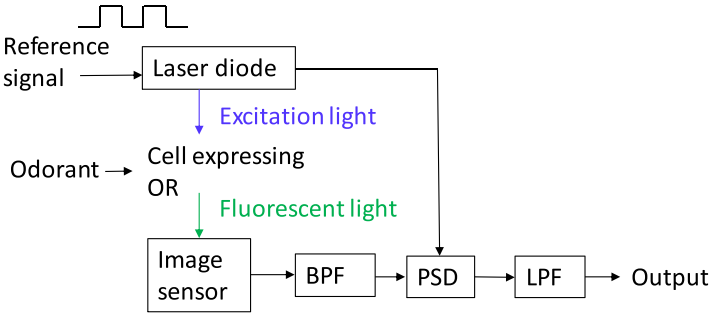

Fig. 4. (Color online) Principle of the lock-in technique.

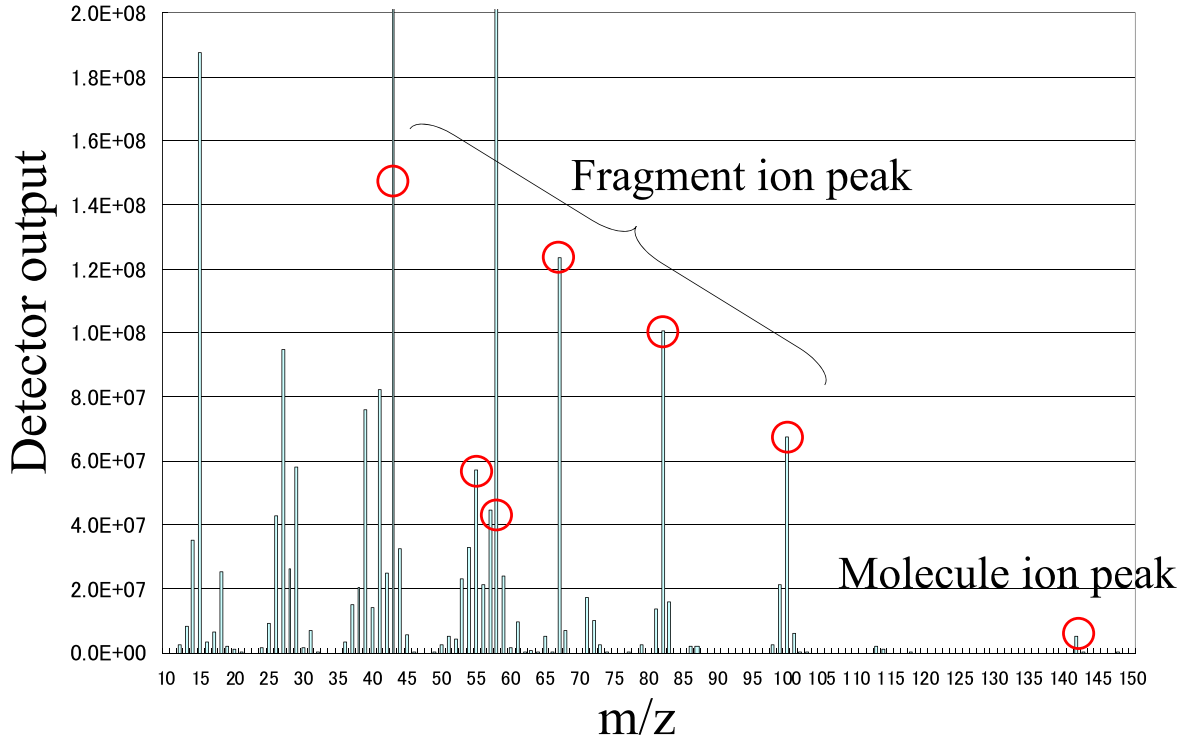

Fig. 5. (Color online) Example of a mass spectrum (trans-2-hexenyl acetate) using the electron ionization method.

classify odors. When the cells were exposed to the odor several times, the sensor response gradually decreased, probably due to photo breaching. However, two types of moldy smell, such as geosmin and 1-octen-3-ol, were separated in spite of large fluctuations of the sensor responses just using Principal Component Analysis (PCA).[24] The pattern separation was achieved because of excellent sensor selectivities originating from OR characteristics.

3. A deep learning technique to predict odor impression

3.1. Prediction of odor impression using a neural network with two autoencoders

The odor impression obtained from sensory testing was predicted using mass spectrum data and a deep learning technique.[18] The mass spectrum is used here since a large-scale database is available and it has large dimensional data including plenty of information. An example mass spectrum is shown in Fig. 5. In addition to a molecular ion peak, many fragment ion peaks appear as m/z (mass to charge ratio) changes. Since each m/z works as a sensor element in a sensor array, a few hundred dimensional data are available. We chose the range of m/z from 51 to 262 since an m/z lower than 51 includes the peaks of solvent and odorless compounds and the region of m/z above 262 is very sparse.

Sensory test data based upon the semantic differential method, and conducted by Dravnieks, were used here.[25] An example of sensory test data is illustrated in Fig. 6. In that database, 144 descriptors including adjectives such as floral, sweet, bitter etc. have scores from zero to five for 121 odorants. Although we would like to have a database that is as large as possible, it includes at most 100 or 200 samples in the field of olfaction. It is difficult to obtain many sensory results under the same conditions when we perform sensory tests of olfaction, whereas a huge quantity of data is available for the visual or auditory senses. The problem of the data size is a critical issue to be solved in olfaction.

A schematic diagram of odor-impression prediction is illustrated in Fig. 7. Its strategy was inspired by the PLS (Partial Least Squares) method.[26] PLS is currently a gold standard of regression, which is known to be superior to multiple linear regression,[27] in the field of chemometrics.[28] In both independent-variable space and dependent-variable space, their features are separately extracted. Then, regression

Sample A

Descriptor	Score
Floral	4
Sweet	3
Bitter	1
⋮	

Sample B

Descriptor	Score
Floral	2
Sweet	1
Bitter	5
⋮	

Fig. 6. Example of sensory data.

between features of independent variables and those of dependent variables are performed. These feature extractions suppress the influence of noise effectively.

The deep learning neural network used here includes two five-layer autoencoders,[29,30] one for input data space (mass spectrum) and the other for output space (sensory data), to extract the features as shown in Fig. 7. Then, the feature vector of the mass spectrum was mapped onto the feature vector of sensory data using a five-layer perceptron (Multi-Layer Perceptron, MLP).[31] The sensory data were reconstructed using the autoencoder at the output space.

The autoencoder is one kind of MLP, where the number of input-layer neurons is the same as that of output-layer neurons. Since the target signals of output-layer neurons are the same as those of inputs, it works as an unsupervised neural network even if the MLP is a supervised one. Since the number of neurons in the middle hidden layer is kept small, the feature of the input data appears on that layer.

The structure in Fig. 7 is similar to that of the PLS. However, we can perform nonlinear feature extraction and nonlinear mapping in our proposed method even though PLS is a linear technique.

The dimensions of feature vectors for mass spectrum and sensory data were 45 and 30, respectively, after optimization. The structure of the five-layer autoencoders for mass spectrum and sensory data were 212-85-45-85-212 and 144-65-30-65-144, respectively. The structure of the MLP between two autoencoders was 45-50-55-50-30. An L1-norm regularization term was added to the cost function to suppress overlearning. In the prediction after training, a nine-layer neural network was used in total.

First, the reconstruction error of the autoencoder was evaluated. Figure 8 shows the relationship between the feature-vector dimension of the mass spectrum and the reconstruction error for both autoencoder and PCA. PCA is a typical linear technique of feature extraction. It was found from the figure that the reconstruction error of the autoencoder was much smaller than that of PCA due to its nonlinear characteristic.

The results of odor-impression prediction are shown in Fig. 9(a) and the results obtained using PLS are shown in Fig. 9(b). The number of latent variables in PLS was 45 after optimization. Both data matrices, the sensory evaluation data and the mass spectra data, were randomly separated into six subsets, five of which (100 samples) were used for training the whole model and the hold-out set (21 samples) was used for evaluation of the generalization error. We repeat this k-fold cross validation 10 times for different random splittings. The true values in the figures come from the database.[25] The plots approach the diagonal line if the accuracy is high. It was found that the correlation coefficient of the proposed method ($\cong$0.76) was higher than that of a conventional method such as PLS ($\cong$0.61). Further improvement is expected if more data are available.

Moreover, there are several tiny peaks at the high m/z region of the mass spectrum. Although the contributions of those small peaks to the whole mass spectrum tend to be ignored, those peaks can contribute to the sensory result since a compound with large molecular weight is detectable even in low concentrations. Thus, a cost function such as the Itakura–Saito divergence to detect small peak change is useful.[19] The topic of predicting odor impression is becoming a hot topic and publications on the subject are increasing.[32]

3.2. Autoencoder with Itakura–Saito cost function

Thus, we studied an autoencoder based upon IS cost function. IS divergence has previously been used for non-negative matrix factorization for music data.[33] The dynamic range of mass spectrum signals is large enough to have a 3–6 order of magnitude. The autoencoder with IS cost function is compared with those of the Euclidean and cross entropy distances to show the improvement in reproduction of small values in a target dataset.

Three cost functions are expressed in the following equations,

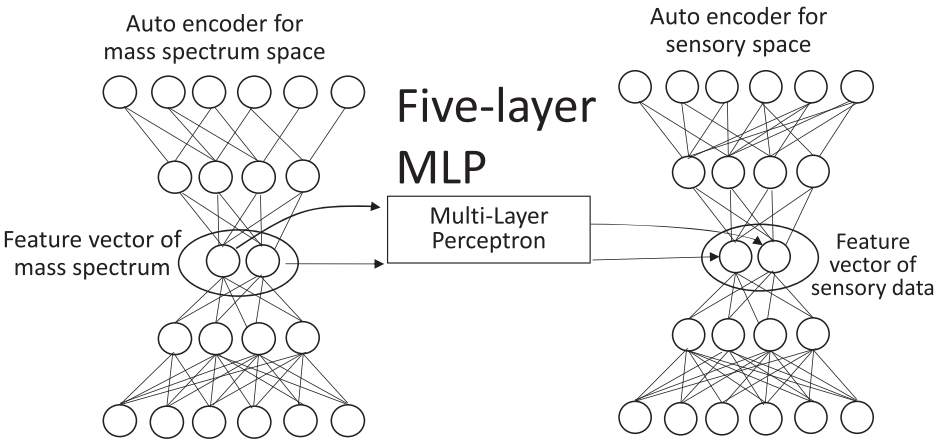

Fig. 7. Schematic diagram of odor-impression prediction.

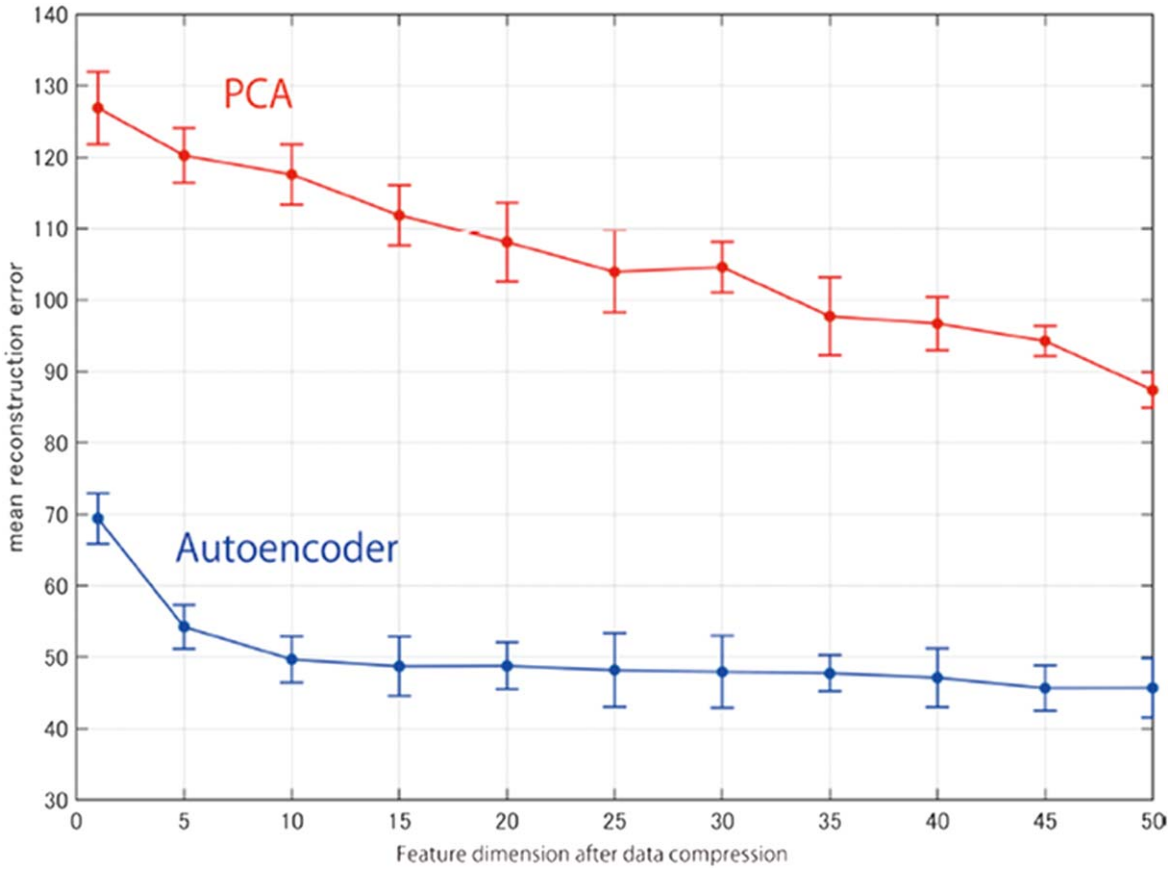

Fig. 8. (Color online) Reconstruction error of the mass spectrum with respect to the feature dimension number.[18)]

$$E_{\mathrm{IS}} = \frac{y}{f(z)} - \ln\frac{y}{f(z)} - 1, \tag{1}$$

$$E_{\mathrm{MSE}} = \frac{1}{2}(y - f(z))^2\,, \tag{2}$$

and

$$E_{\mathrm{CE}} = y\ln f(z) + (1 - y). \tag{3}$$

E_{IS} is the IS-based cost function, E_{MSE} is the MSE (Mean Squared Error) and E_{CE} is the CE (Cross Entropy). y is a target signal, an element of the original mass spectrum given to an autoencoder, and $f(z)$ is an output of sigmoid function on input value z. This is a one-dimensional case for simplicity. It is calculated for each dimension, followed by summing all dimensions for the multi-dimensional case. The IS-based cost function changes drastically when values are near zero, as compared with the other cost functions.

Since an autoencoder is a kind of MLP, a back-propagation algorithm was used to train it. Gradients of a cost function with respect to the weights and the biases in the network are calculated in the back-propagation algorithm. The weights and biases are repeatedly updated according to the gradients so that the cost function can be minimized. Thus, the method to minimize the cost function is called the gradient descent method. The gradients should be iteratively calculated layer by layer so that the gradient descent method can optimize the weights in the entire network. The derivation of the equation for updating weights and biases is described in Ref. 19.

To compare the performances of neural networks based on three different cost functions, a preliminary experiment on a simple perceptron with a single neuron, shown in Fig. 10, was performed. This is the simplest network to observe the fundamental capability of cost functions. The perceptron was iteratively trained to output a small value close to zero when the input was one. The learning speed and errors of perceptrons with different cost functions were evaluated in this experiment.

The target signal was chosen from 0.1 to 0.0001 and the error between the target signal and the output after 300 learning epochs was calculated in order to understand how the cost function influenced the error. Two sets of initial weight and bias were used since the training speed is highly sensitive to the initial values of weights and biases.

Figure 11 shows the results at each condition. The smaller a target value was, the larger the errors of the perceptron with MSE and CE cost functions became, in both cases. On the other hand, the output of the perceptron with IS cost function was much closer to the target value than other perceptrons' outputs at both initial conditions. In particular, the MSE autoencoder shows slower convergence.

These undesirable errors of MSE and CE are caused by a slowdown in learning speed. When both the output and the target values are small, this slowdown is inevitable as the subtraction of two values, $y - f(z)$, governs perceptrons' gradients. However, when using IS divergence as a cost function, the ratio of y to $f(z)$ governs the gradient when

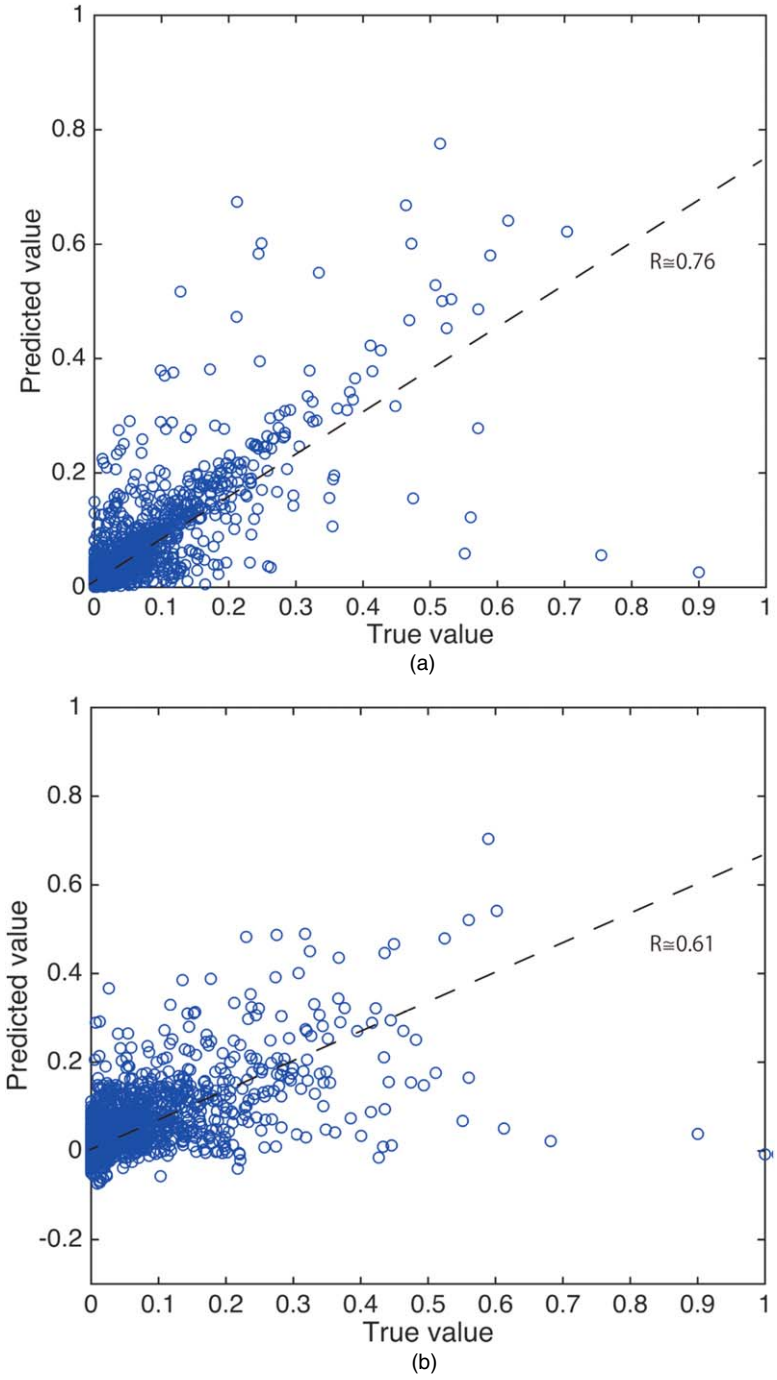

Fig. 9. (Color online) Result of odor-impression prediction. (a) Proposed method and (b) PLS method.[18] Score is normalized between zero and one.

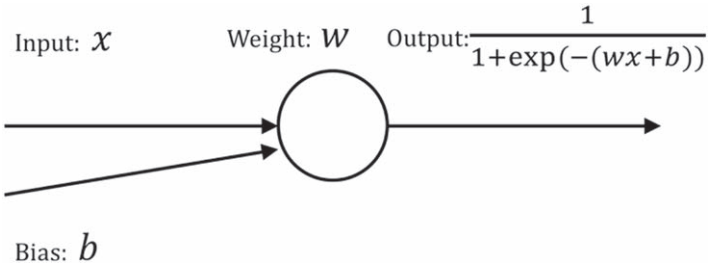

Fig. 10. Structure of a simple perceptron with a single neuron. (Copyright Elsevier 2017 with permission).[19]

both values are small. This resulted in the improvement of learning speed of the perceptron. Therefore, as we expected, the IS-based perceptron has a much higher accuracy of outputting small values close to zero.

The results of earlier experiments showed that a perceptron with IS divergence as its cost function is promising. Our next target is to expand this perceptron to a deep autoencoder so that it can be applied to a large-scale dataset such as a set of

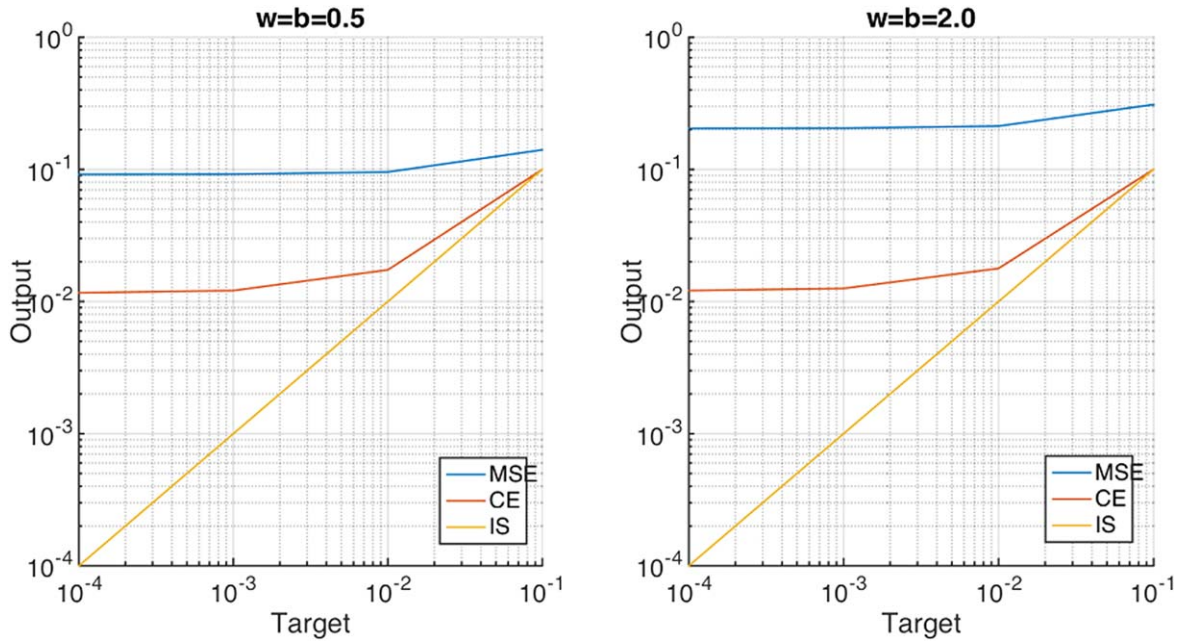

Fig. 11. (Color online) Plots of output value versus target value (from 0.1 to 0.0001) after 300 training epochs. Each autoencoder is trained under the initial condition of $w = b = 0.5$ (left) and $w = b = 2.0$ (right). (Copyright Elsevier 2017 with permission).[19)]

mass spectra of chemical compounds. The five-layer sand-glass-shaped autoencoder was used for the experiment of dimensionality reduction of mass spectra.

Mass spectrum data used in this study were obtained from the Chemistry WebBook provided by the National Institute of Standards and Technology.[34)] This database has more than 100 000 mass spectra acquired using the electron ionization method with an energy of 70 eV. Each peak in the entire database was normalized to a range between zero and one, and then 5000 samples were randomly picked and divided into two sets, a training set of 3000 samples and a testing set of 2000 samples. Although most of the mass spectra are sparse and include a lot of zero values, all the zero values in the dataset were replaced with a small value (=0.01). Since a chemical compound with a molecular weight higher than 300 is typically non-volatile, intensities with m/z values between 1–300 were used in the experiment.

An average error relative to target signal was calculated after the computational experiment. Relative error was calculated as:

$$\text{Relative Error} = \frac{1}{s}\sum\frac{|y - f(z)|}{y}[\%], \tag{4}$$

where s is the number of samples. To evaluate the reproduction capabilities of small values, relative error was calculated for two groups of peaks which are split from the viewpoint of dynamic range, namely peaks of large value (from 3% to 100% of full dynamic range) and peaks of small value (below 3% of full dynamic range). The maximum value in the dataset was normalized to have unity value. The peaks between 0.03–1.0 correspond to the former group and the peaks between 0.01–0.03 correspond to the latter group. The error ratio of each peak group is summarized in Table I. The autoencoder with IS cost function achieved the smallest relative error for the group of small peaks, whereas the CE autoencoder did for the group of large peaks. When the threshold value was changed from 3% to 1% or 5%, similar results were obtained. Thus, IS divergence is found to be the most appropriate cost function on this dataset in terms of the reproduction accuracy for small target inputs. The error relative to target signal in Table I was calculated for the validation dataset. The experiment on the validation dataset reveals that the improvement effect on small values using IS cost function was not disturbed by over-fitting, and that the IS-based autoencoder was well generalized in unknown samples.

Table I. Averaged error relative to target signal in the test dataset. (Copyright Elsevier 2017 with permission).[19)]

Cost function	Relative error for large signals [%]	Relative error for small signals [%]	Relative error for all signals [%]
MSE	36.86	70.60	68.52
CE	35.43	57.44	56.10
IS	41.55	38.01	38.23

Thus, the IS-based autoencoder showed a significant improvement in its approximation capabilities for small values, which resulted in improvement in the approximation of the entire mass spectra dataset, while slightly sacrificing its approximation capabilities for large values.

4. Conclusion

The odor biosensor and the deep learning for odor-impression prediction were separately explained here, although it will be possible to combine them in the future. In the former part, it is possible to classify odors using randomly distributed cells expressing different ORs. The study of odor biosensors should be further extended so that more complicated information such as mixture composition can be extracted. The latter part is the prediction of odor impression using mass spectra and deep learning. Although there is still room for improvement in prediction accuracy, it might be enhanced if we obtain huge quantities of data, especially sensory data. Moreover, the relationship between odor descriptors is obtained from natural language processing.[35)]

Furthermore, it is feasible to expand these techniques to odor reproduction using odor components.[36,37] Other work such as odor recording and olfactory displays described elsewhere[38,39] can be introduced at the next opportunity.

Acknowledgments

The author wishes to thank Prof. Ryohei Kanzaki and Dr. Hidefumi Mitsuno for providing Sf21 cells expressing olfactory receptors. This work was partially supported by JSPS KAKENHI Grant Numbers JP18H03773 and JST Mirai Program Grant Number JPMJMI17DD.

ORCID iDs

Takamichi Nakamoto https://orcid.org/0000-0002-0599-226X

1) G. M. Shepherd, *Neurobiology* (Oxford University Press, Oxford, 1988), p. 222.
2) K. Persaud and G. Dodd, "Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose," Nature **299**, 352 (1982).
3) T. Nakamoto and T. Moriizumi, "Odor sensor using quartz-resonator array and neural-network pattern recognition," IEEE Ultrason. Symp., p. 613 (1988).
4) T. C. Pearce, S. S. Schiffman, H. T. Nagle, and J. Gardner, *Handbook of Machine Olfaction* (Wiley, Weinheim, 2003).
5) M. Kaneyasu, A. Ikegami, H. Arima, and S. Iwanaga, "Smell identification using thick-film hybrid gas sensor," IEEE Trans. Compon. Hybrids Manuf. Technol. **10**, 267 (1987).
6) H. Sundgren, F. Winquist, and I. Lundstrom, "Artifical neural networks and statistical pattern recognition improve MOSFET gas sensor array calibration," Tech. Dig. Transducers **91**, 574 (1991).
7) M. E. H. Amrani, R. M. Dowdeswell, P. A. Payne, and K. C. Persaud, "An intelligent gas sensing system," Sens. Actuators B **44**, 512 (1997).
8) J. W. Grate, "Acoustic wave microsensor arrays for vapor sensing," Chem. Rev. **100**, 2627 (2000).
9) Y. K. Yoo, M. Chae, J. Y. Kang, T. S. Kim, K. S. Hwang, and J. H. Lee, "Multifunctionalized cantilever systems for electronic nose applications," Anal. Chem. **84**, 8240 (2012).
10) T. Nakamoto, A. Fukuda, and T. Moriizumi, "Perfume and flavour identification by odour-sensing system using quartz-resonator sensor array and neural-network pattern recognition," Sens. Actuators B **10**, 85 (1993).
11) N. A. Rakow and K. S. Suslick, "A colorimetric sensor array for odour visualization," Nature **406**, 710 (2000).
12) J. R. Stetter, P. C. Jurs, and S. L. Rose, "Detection of hazardous gases and vapors: pattern recognition analysis of data from an electrochemical sensor array," Anal. Chem. **58**, 860 (1986).
13) M. Bernabei, K. C. Persaud, S. Pantalei, E. Zampetti, and R. Beccherelli, "Large-scale chemical sensor array testing biological olfaction concepts," IEEE Sens. J. **12**, 3174 (2012).
14) J. H. Sung, H. J. Ko, and T. H. Park, "Piezoelectric biosensor using olfactory receptor protein expressed in *Escherichia coli*," Biosens. Bioelectron. **21**, 1981 (2006).
15) K. Sato and S. Takeuchi, "Chemical vapor detection using a reconstituted insect olfactory receptor complex," Angew. Chem. Int. Ed. **53**, 11798 (2014).
16) Y. Sukekawa, T. Mujiono, T. Nakamoto, H. Mitsuno, Y. Nakajima, R. Kanzaki, and N. Misawa, Electron. Commun. Jpn. **100**, 41 (2017).
17) G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science **313**, 504 (2006).
18) Y. Nozaki and T. Nakamoto, "Odor impression prediction from mass spectra," PLoS ONE **11**, e0157030 (2016).
19) Y. Nozaki and T. Nakamoto, "Itakura-Saito distance based autoencoder for dimensionality reduction of mass spectra," Chemometrics Intell. Lab. Syst. **167**, 63 (2017).
20) T. Nakamoto, Abstract of SSDM 2018, K-7-01, 2018.
21) H. Mitsuno, T. Sakurai, S. Namiki, H. Mitsuhashi, and R. Kanzaki, "Novel cell-based odorant sensor elements based on insect odorant receptors," Biosens. Bioelectron. **65**, 287 (2015).
22) T. Mujiono, Y. Sukekawa, T. Nakamoto, H. Mitsuno, M. Temtanasombat, R. Kanzaki, and N. Misawa, "Sensitivity improvement by applying lock-in technique to fluorescent instrumentation for cell-based odor sensor," Sens. Mater. **29**, 65 (2017).
23) T. Mujiono, Y. Sukekawa, T. Nakamoto, H. Mitsuno, Y. Nakajima, R. Kanzaki, and N. Misawa, "Lock-in measurement technique in fluorescent instrumentation system for cell-based odor sensor," IEEJ Trans. Sens. Micromach. **136**, 83 (2016).
24) Y. Sukekawa and T. Nakamoto, "Odor biosensor system based on image lock-in measurement for odorant discrimination," Electron Comm. Jpn. **102**, 57–64 (2019).
25) A. Dravnieks, *Atlas of Odor Character Profiles* (ASTM, Philadelphia, PA, 1985).
26) P. Geladi and B. R. Kowalski, Anal. Chim. Acta **185**, 1 (1986).
27) W. R. Dillon and M. Goldstein, *Multivariate Data Analysis* (Wiley, New York, 1984), p. 209.
28) M. A Sharaf, D. L. Illman, and B. R Kowalski, *Chemometrics* (Wiley, New York, 1986).
29) R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification* (Wiley, New York, 2001), p. 568.
30) I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, 2016), p. 493.
31) G. Dougherty, *Pattern Recognition and Classification* (Springer, Berlin, 2013), p. 127.
32) A. Keller et al., Science **355**, 820 (2017).
33) C. Févotte, N. Bertin, and J. Durrieu, "Nonnegative matrix factorization with the Itakura-Saito divergence: with application to music analysis," Neural Comput. **21**, 793 (2009).
34) NIST Chemistry WebBook, http://webbook.nist.gov/chemistry/.
35) Y. Nozaki and T. Nakamoto, "Predictive modeling for odor character of a chemical using machine learning combined with natural language processing," PLoS ONE **13**, e0198475 (2018).
36) T. Nakamoto, M. Ohono, and Y. Nihei, "Odor approximation using mass spectrometry," IEEE Sens. J. **12**, 3225 (2012).
37) T. Nakamoto and Y. Nihei, "Improvement method of odor approximation using mass spectrometry," IEEE Sens. J. **13**, 4305 (2013).
38) T. Nakamoto (ed.) *Essentials of Machine Olfaction and Taste* (Wiley, New York, 2016).
39) T. Nakamoto (ed.) *Human Olfactory Displays and Interfaces* (IGI-Global, Hershey, PA, 2013).

Takamichi Nakamoto received his B.Eng. and M.Eng. degrees in 1982 and 1984, respectively, and his PhD degree in electrical and electronic engineering from the Tokyo Institute of Technology, Tokyo, Japan. He worked for Hitachi from 1984 to 1987. In 1987, he joined the Tokyo Institute of Technology as a Research Associate. In 1993, he became an Associate Professor with the Department of Electrical and Electronics Engineering, Tokyo Institute of Technology. From 1996 to 1997, he was a Visiting Scientist at Pacific Northwest Laboratories, Richland, WA, USA. He is currently a Professor with the Institute of Innovative Research, Tokyo Institute of Technology.

