

Odor Clustering Based on Molecular Parameter for Odor Sensing

Masahiro Imahashi* and Kenshi Hayashi

Department of Electronics, Graduate School of Information Science and Electrical Engineering,
Kyushu University, 744 Motoooka, Nishi-ku, Fukuoka 819-0395, Japan

(Received October 1, 2013; accepted November 21, 2013)

Key words: odor map, molecular parameter, pattern recognition, odor clustering, multidimensional scaling

Odor sensors can benefit various areas of human activity and have been increasingly studied. For developing odor sensors, comprehensive detection of numerous volatile molecules is necessary. These advanced odor measurements might be accomplished by inspiring technology based on the bio-olfactory system. This system recognizes and discriminates odors by activity patterns, which are formed based on odor information of odorants extracted from olfactory receptors (ORs). Hence, odorants are appropriately categorized into clusters with different molecular features. The odor clustering close to biological olfaction can also be applied to the sensor systems. In this study, odor map images of rats investigated in biological studies were analyzed by principal component analysis (PCA) to clarify odor clustering features of olfaction. The definition of odor cluster and extraction of geometric features of odor maps were examined based on the primary components and factor loadings. Then, key parameters expressing clusters and measurable in sensor technology were successfully explored by evaluating the correlation between principal components and molecular parameters calculated using the molecular modeling software. Finally, artificial odor maps were reconstructed based on the defined odor clustering map, and the similarity between odor maps of rats was confirmed.

1. Introduction

Recently, the needs for sensing of dangerous or hazardous volatile organic components (VOCs) and the simultaneous monitoring of various chemicals in air have been increasing.^(1–3) Sensor devices using an unspecific sensor array, *e.g.*, E-nose technology, enable the detection and classification of some VOCs.^(4–7) However, since it is said that there exist hundreds of thousands of types of odorant that humans can perceive,⁽⁸⁾ it is very difficult to measure numerous chemicals and odors consisting of various chemicals using cross-selective sensor arrays. Therefore, sensor systems that can comprehensively detect and simultaneously discriminate odors are required.

*Corresponding author: e-mail: imahashi@o.ed.kyushu-u.ac.jp

Here, biological olfaction performs these comprehensive measurements and discrimination of the differences in odors. It is thought that humans can detect volatiles in ppt - ppm levels and express the differences among odors.⁽⁹⁾ Therefore, high performance in odor measurements might be accomplished using a sensor technology that imitates the recognition mechanism using molecular parameters of odorants as bio-olfactory recognition processes. In the olfactory mechanism, when mammals smell odors, olfactory receptors (ORs) trap constituent odorants and recognize molecular properties, *e.g.*, complex shapes and functional groups of molecules.⁽¹⁰⁾ Then, the active signals of ORs are separately transmitted to glomeruli on the olfactory bulb (OB)^(11,12) which is the first terminal of olfaction. Finally, the activity pattern of the odor (odor map) is formed on the surface of OB based on OR signals.⁽¹³⁻¹⁵⁾ Thus, the molecular information of odorants is integrated into odor maps.

In addition, it is found that the areas in odor maps activated by each odorant are definitely located, *i.e.*, the odor map is categorized into clusters with different molecular features. For instance, glomeruli activated by fatty acids belong to cluster A.^(8,16,17) This geometrical categorization is called "odor clustering", where the positional relation among clusters is important. Furthermore, humans cannot perceive every constituent of odors separately. Odor maps formed based on OR signals become the first internal expressions of odor in the brain. In fact, biological olfaction carries out the compression of numerous numbers of odor information, and then discriminates the differences in odors on the basis of these clustering patterns.

Advanced odor sensing could be realized using sensor devices that can construct an odor clustering map of biological olfaction. Therefore, with the aim of sensor applications, the definition of odor clustering and the extraction of key parameters essential to bio-olfactory odor measurement were examined by analyzing biological images that represent the activity patterns formed on OB of rats. Then, artificial odor maps were constructed to confirm the accuracy of odor clustering and visualize the odor information. The present analysis makes it possible to relate molecular information of odorants and molecular parameters that can be detected using physicochemical principals.

2. Methods

2.1 Analysis of odor maps

As mentioned in the previous section, the molecular profiles of odorants and clustering features acquired for measuring odors were projected into odor maps. We analyzed odor mapping images of rats to determine clustering processes of biological olfaction. These images have been developed using uptake of radiolabeled 2-deoxyglucose in the Leon laboratory⁽¹⁸⁻²⁰⁾ and are provided publicly on their website (<http://gara.bio.uci.edu/>) as grayscale images of rats equivalent to 321 types of odorant.

In this study, the extraction of fingerprints of clusters and dimensionality reduction were attempted using principal component analysis (PCA). Odor maps of rats were analyzed as follows. First, 197×357 pixel matrices were prepared from 321 types of grayscale image. Second, after the basis vectors have been calculated from 321 types of

matrix with PCA, pixels not attributed to the activity patterns were removed. PCA was calculated on the autoscaled data set. Third, treated matrices were repeatedly analyzed by PCA. The principal component (PC n) scores and factor loadings of rat images were calculated. Finally, concerning each PC n , pixels with smaller absolute valued factor loading were also removed.

2.2 Correlation among principal components and molecular parameters

To explore key parameters essential to odor clustering, the correlation between principal components produced in § 2.1 and various molecular parameters of odorants was evaluated. 76 types of molecular parameter of 321 types of odorant corresponding to grayscale images of rats were calculated using computational packages, *e.g.*, MOPAC and GAMESS, of chemBio3D software (CambridgeSoft). The calculated parameters are shown in Table 1. In this study, Pearson's product-moment correlation coefficients were computed to confirm the correlation between PC n and each molecular parameter.

2.3 Construction of artificial odor maps

It was assumed that activity regions of each cluster in odor maps become "ellipse" shapes. Artificial odor maps were composed of ellipses representing clusters and the activity "intensity". Two coordinates x and y of maps coincide with the odor maps of rats developed in the Leon laboratory.^(18–20) Because the regions around active parts are gradually weakening in odor maps, the activity pattern in each cluster could be assumed to be decreasing from the center of the ellipse in accordance with Gaussian function. Finally, grayscale images of artificial maps were created on the basis of the position of

Table 1
Molecular parameters calculated using computational properties of chemBio3D software.

	Molecular parameter		
Balaban index	DIPOLE-BASED-GAM ZZZZ	Hydrophilic-lipophilic balance (HLB)	Polar surface area
Boiling point	Elemental analysis (C)	Ideal gas thermal capacity	Potential energy
Cluster count	Elemental analysis (H)	Ionization potential	Principal moment
Connolly accessible area	Elemental analysis (N)	Kinetic energy	Radius
Connolly molecular area	Elemental analysis (O)	Partition coefficient (Log P)	Shape attribute
Connolly solvent excluded volume	Elemental analysis (S)	m/z	Shape coefficient
Core-core repulsion	ENERGY-BASED-ALPHA ZZ	Melting point	Straight carbon number
COSMO area	ENERGY-BASED-BETA ZZZ	Mol refractivity	Sum of valence degrees
COSMO volume	ENERGY-BASED-DIPOLE Z	Molecular depth	Surface area
Critical pressure	ENERGY-BASED-GAM ZZZZ	Molecular length	Topological diameter
Critical temperature	Estimation of Henry's constant (H)	Molecular topological index	Total dipole
Critical volume	Exact mass	Molecular volume	Total energy
Dipole moment (MOPAC)	Formal charge	Molecular weight	Total valence connectivity
DIPOLE-BASED-ALPHA XZ	Freezing point	Molecular width	Vapor pressure
DIPOLE-BASED-ALPHA YZ	Gibbs free energy	MR	Water solubility
DIPOLE-BASED-ALPHA ZZ	H bond acceptor	Normal boiling point	Wiener index
DIPOLE-BASED-BETA XZZ	H bond donor	Number of rotatable bonds	
DIPOLE-BASED-BETA YZZ	Hansen polarity	Ovality	
DIPOLE-BASED-BETA ZZZ	Heat of formation	Percent hydrophilic surface	
DIPOLE-BASED-DIPOLE Z	Henry's law constant	Acid dissociation constant (pKa)	

each cluster and activity pattern. These grayscale images were colored using ImageJ software (NIH).

Artificial maps produced in this study were transformed to odor map images of rats. Therefore, the intensities of activity patterns in each cluster were calculated from the average brightness value of the cluster ellipse in rat maps. Here, if the sensor responses corresponding to each cluster can be obtained, the construction of artificial odor maps classified into clusters becomes possible. Hence, the comprehensive detection and appropriate categorization of odors might be realized.

3. Results and Discussion

3.1 Definition of odor clustering map

Principal components were calculated from 321 types of grayscale image of rats with PCA. As a result, the cumulative contribution from PC1 to PC80 has reached 80%. Concerning each principal component, factor loadings responding to all pixels were determined. Therefore, odor images were distributed depending on both positive and negative values of loadings. However, because most factor loadings were smaller in PC7 to PC80, most corresponding pixels were removed. Therefore, areas allotted in odor maps became narrow. In this study, important principal components from 1 to 6 were focused on. Thus, these important principal components colored as the contour scale are shown in Fig. 1. These figures suggested that the odor images could be divided into 9

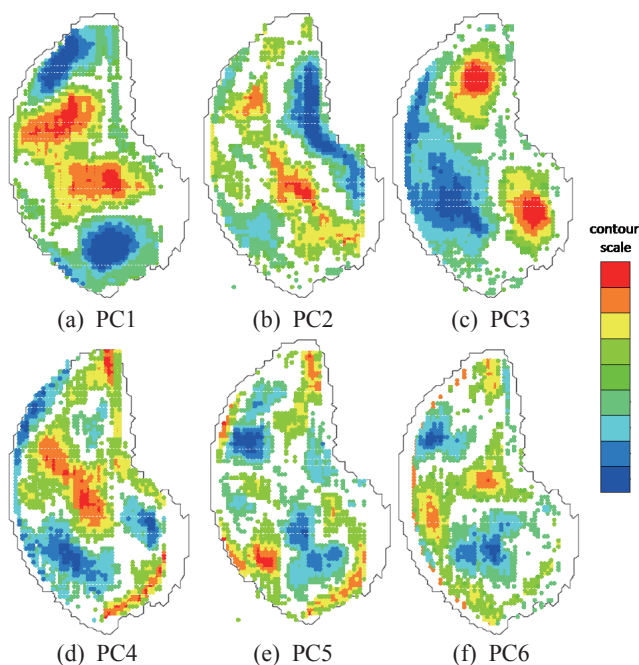


Fig. 1. (Color online) Odor map images distributed by factor loadings of PC1 to PC6, which were colored as the contour scale.

independent areas. To confirm whether these areas are equal to odor clusters, principal scores were plotted in Fig. 2.

Figure 2 shows that odorants with cyclic and straight structures were grouped in higher positive and negative areas of PC1. Specifically, odorants with more than nine carbon chain length were scattered in positive areas. Moreover, in higher positive and negative areas of other components, odorants with the same molecular features, *e.g.*, hydrocarbon, fatty acid, aldehyde, and phenol, were also gathered. These results represent that nine activity regions divided with PCA were classified on the basis of information on molecular size and functional groups.

Here, individual ORs recognize odorants with certain several consecutive carbon chain lengths and certain functional groups, and then biological olfaction forms odor maps using different combinations of active signals from ORs.⁽¹⁰⁾

It is considered that the odor maps are also classified on the basis of molecular features of odorants. The olfactory mechanism was consistent with PCA experiment results that activity patterns were assorted on the basis of the carbon-chain number and

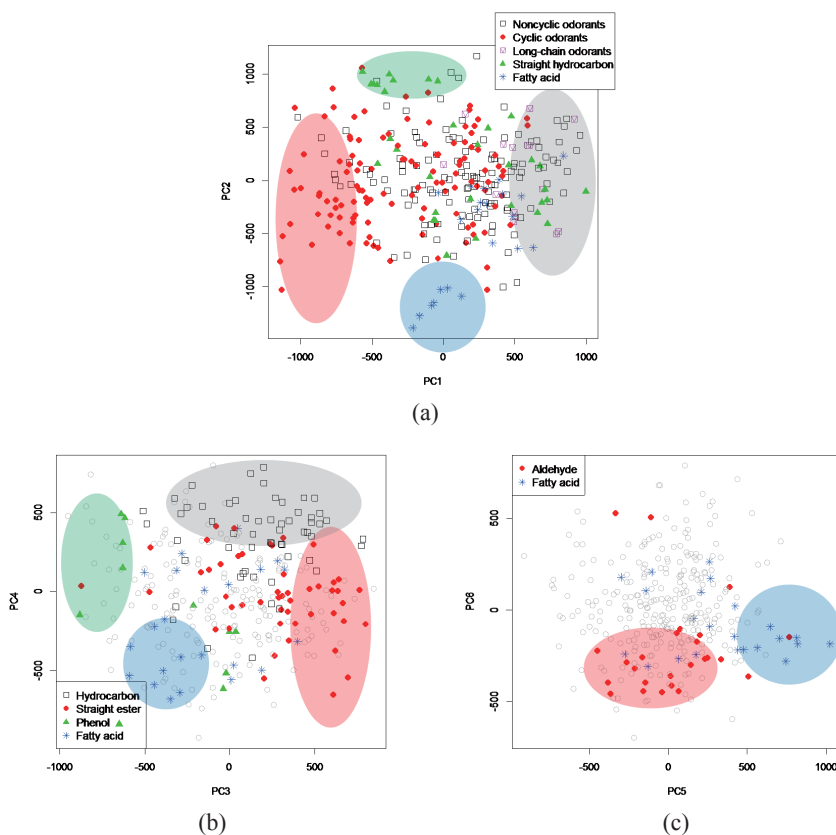


Fig. 2. (Color online) Scatter plots of (a) PC1 and PC2, (b) PC3 and PC4, and (c) PC5 and PC6 of the PCA model calculated using 321 types of grayscale image of rats.

functional groups of odorants. Hence, these categories are potentially equivalent to odor clustering of biological olfaction. The nine classified clusters and the common characteristics of odorants activated in each part are described in Fig. 3. These nine clusters were obtained through a combination of areas corresponding to the six described principal components. This odor clustering map has no discrepancy with the results of other studies. For instance, the cluster A area was activated by fatty acids.^(8,17) A variety of odorants can be categorized into the corresponding clusters, by viewing activity patterns of the odor maps of rats.

3.2 Extraction of key parameters describing odor maps

It was shown that odor measurements were realized by discriminating functional groups of constituent odorants. However, it is difficult to measure only the partial structure of molecules. To apply odor clustering to the sensor technology, a measurable molecular parameter corresponding to each cluster needs to be explored. For the extraction of such key parameters, correlation coefficients between PC n and molecular parameters obtained using the molecular modeling software were calculated. The correlation results are shown in Table 2. In Table 2, correlated parameters and maximum correlation coefficients are described. As a result, the parameters correlated with positive and negative principal components could be explored. Because the number of rotatable bonds and molecular size have more associations with PC1, higher or lower regions [Fig. 1(a)] were potentially activated by odorants with benzene rings or larger carbon chain. Moreover, it was found that the polarity information, *e.g.*, acid dissociation constant (pKa) and partition coefficient ($\log P$), can become key information parameters. It is quite

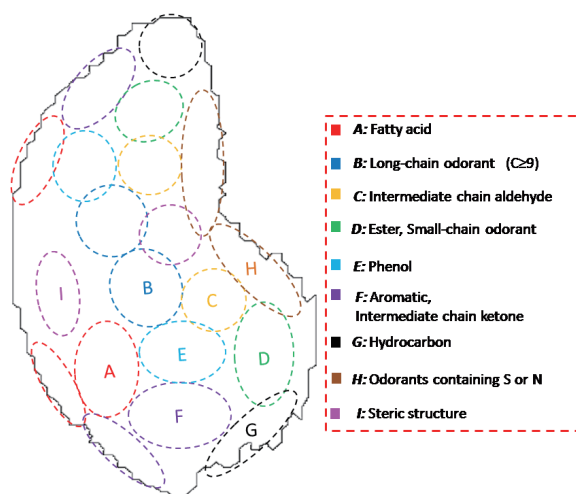


Fig. 3. (Color online) PCA-combined odor clustering map classified by functional groups of odorants.

Table 2

Correlation results between principal component scores and molecular parameters. Correlated parameters and maximum correlation coefficients are described.

PC _n	PC1		PC2	
	+	-	+	-
Correlated parameters	Number of rotatable bonds	Molecular length	Water solubility	Elemental analysis (I, S)
Maximum correlation coefficient	0.56	0.56	-0.72	0.42

PC _n	PC3		PC4	
	+	-	+	-
Correlated parameters	Ester structure	pKa Critical temperature Boiling point	Percent hydrophilic surface Straight carbon number Log <i>P</i> HLB	Polar surface area Log <i>P</i> Partition coefficient
Maximum correlation coefficient		-0.62	0.63	-0.69

PC _n	PC5		PC6	
	+	-	+	-
Correlated parameters	pKa	Vapor pressure		pKa
Maximum correlation Coefficient	0.35	0.51		0.42

reasonable that these parameters were chosen because it is said that they are important factors in receptor science. However, not all the parameters explaining clusters or representing the positional relation between clusters were searched.

The geometric progressions in odor maps of rats correlated with key parameters are described in Fig. 4. Arrows indicate the activity progression by various incremental differences in structure. As shown in Fig. 4, the regions of odor maps are classified by the degree of hydrophobicity and hydrophilicity of volatile chemicals, and activity parts progressively move according to increasing log *P* or hydrophilic-lipophilic balance (HLB) (blue arrow in Fig. 4). Furthermore, chemotopic progressions by pKa and molecular size also occur in hydrophobic and hydrophilic regions. Consequently, odor clustering maps could be roughly constructed using these key parameters, which can be measured by physicochemical methods such as multichannel sensor arrays.

3.3 Imaging of odor information

Because odor maps contain various molecular properties, artificial odor maps enable not only the visualization but also comprehensive qualitative classification of odors. In this study, artificial maps of three types of odorant, butyric acid, decanal,

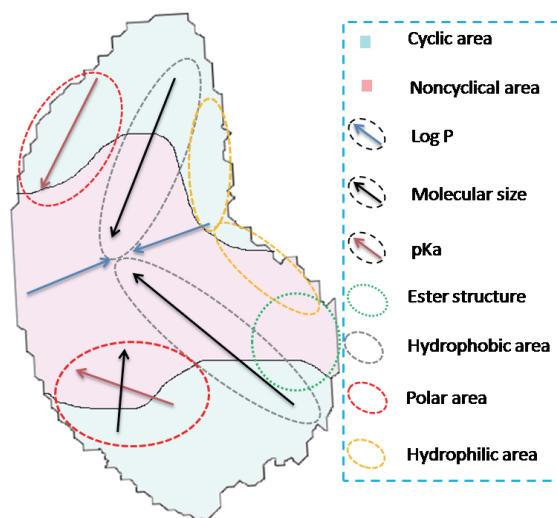


Fig. 4. (Color online) Odor map images and geometric progression described by correlated key parameters with principal component scores.

and acetoxyacetone, which belong to clusters A, B, and D from odor map patterns, respectively, were constructed on the basis of nine ellipses and activity intensities calculated from rat images. For evaluating constructed maps and verifying the accuracy of defined odor clustering, the similarity of individual odor map images was expressed using the multidimensional scaling (MDS) method. MDS can visualize datasets representing the cosine similarity between prepared images, preserving the distance between images as much as possible. Therefore, images with similar patterns can be plotted closely in MDS. In this study, the cosine similarity was calculated between active regions of images.

The similarity between artificial maps and rat maps corresponding to 35 odorants was evaluated by the MDS method. These maps of rats are categorized into different odor clusters. Figure 5 shows an MDS scatter plot. Odorant images of rats classified clearly into the same clusters on the basis of odor clustering maps in Fig. 3 were closely grouped in Fig. 5. This indicates that individual patterns activated by odorants belonging to the same cluster were similar. We considered that the defined odor clustering was close to that of biological olfaction.

In addition, artificial odor maps have high similarity to raw odorant images and could be classified into the same cluster. We concluded that the odor map imaging proposed in this study enables the compression of odor information acquired for recognizing and discriminating odors by the same manner as the olfactory mechanism. Then, the visualization and classification of various odors become possible using odor clustering maps.

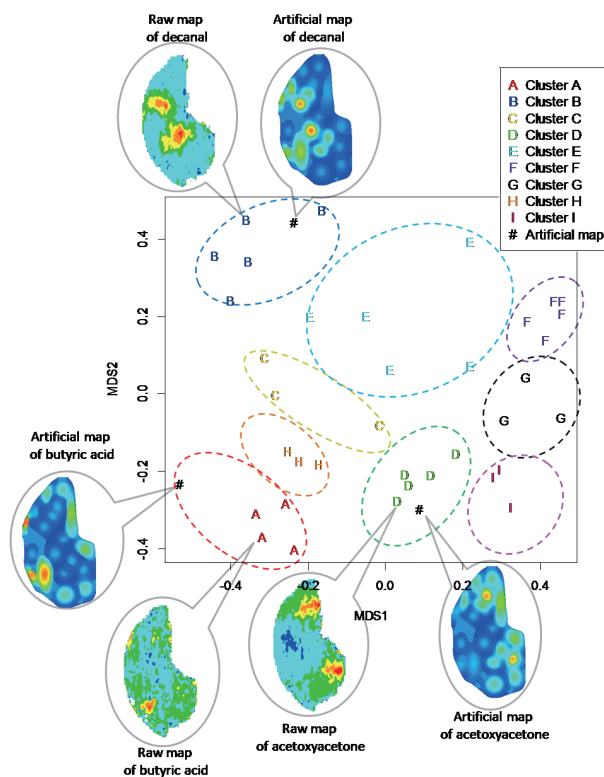


Fig. 5. (Color online) Similarity plot visualized by MDS. Plotted maps correspond to rat images of odorants belonging to clusters and artificial odor maps. Raw maps were cited from public data of the Leon laboratory.

4. Conclusions

We succeeded in defining appropriate odor clustering using odor map images of rats with PCA and exploring important molecular parameters for odor clustering. Then, odor images of rats were successfully transformed and efficiently compressed to artificial odor maps. Thus, odor maps are easily and precisely constructed by measuring these parameters with a sensor array.

In our other works, the sensor system with high molecular recognition ability for extracting molecular parameters from measured volatiles was already realized.^(21,22) That system can detect and measure various odors using a molecularly imprinted filter.^(23,24) Consequently, comprehensive odor sensing is possible using the developed sensor system incorporated with defined odor clustering shown in this study.

Acknowledgements

This work was partially supported by JSPS KAKENHI Grant Number 252501. We are very grateful to Dr. M. Leon and staff members of his laboratory for publicly presenting the odor map images of rats on their website.

References

- 1 H. Nanto, Y. Yokoi, T. Mukai, J. Fujioka, E. Kusano, A. Kinbara and Y. Douguchi: *Mater. Sci. Eng., C* **12** (2000) 43.
- 2 T. Oka, S. Hayashida, Y. Kaneda, M. Takenaga, Y. Tamagawa, S. Tsuji and A. Hatanaka: *Biopsychosoc. Med.* **2** (2008) 1.
- 3 G. Horvath, G. A. Järverud, S. Järverud and I. Horváth: *Integr. Cancer Ther.* **7** (2008) 76.
- 4 T. Pearce, S. Schiffman, H. Nagle and J. Gardner: *Handbook of Machine Olfaction: Electronic Nose Technology* (Wiley-VCH, London, 2003).
- 5 I. Elmi, S. Zampolli, E. Cozzani, F. Mancarella and G. C. Cardinali: *Sens. Actuators, B* **135** (2008) 342.
- 6 A. Bartolazzi, M. Santonico, G. Pennazza, E. Martinelli, R. Paolesse, A. D'Amico and C. Di Natale: *Sens. Actuators, B* **146** (2010) 483.
- 7 K. Arshak, E. Moore, G. M. Lyons, J. Harris and S. Clifford: *Sens. Rev.* **24** (2004) 181.
- 8 K. Mori, Y. K. Takahashi, K. Igarashi and M. Yamaguchi: *Physiol. Rev.* **86** (2006) 409.
- 9 M. Yatagai and M. Kawasaki: *Aroma and environment: Aroma Science Series 21* (Fragrance Journal, Tokyo, 2002) (in Japanese).
- 10 B. Malnic, J. Hirono, T. Sato and L. Buck: *Cell* **96** (1999) 713.
- 11 L. Buck and R. Axel: *Cell* **65** (1991) 175.
- 12 L. Buck: *Nutr. Rev.* **62** (2004) 184.
- 13 S. Korsching: *Curr. Opin. Neurobiol.* **12** (2002) 387.
- 14 E. Soucy, D. Albeanu, A. Fantana, V. Murthy and M. Meister: *Nat. Neurosci.* **12** (2009) 210.
- 15 T. Bozza, A. Vassalli, S. Fuss, J. J. Zhang, B. Weiland, R. Pacifico, P. Feinstein and P. Mombaerts: *Neuron* **61** (2009) 220.
- 16 K. Mori and Y. Yoshihara: *Prog. Neurobiol.* **45** (1995) 585.
- 17 H. Matsumoto, K. Kobayakawa, R. Kobayakawa, T. Tashiro, K. Mori, H. Sakano and K. Mori: *J. Neurophysiol.* **103** (2010) 3490.
- 18 B. Johnson and M. Leon: *J. Comp. Neurol.* **503** (2007) 1.
- 19 B. Johnson and M. Leon: *J. Comp. Neurol.* **422** (2000) 496.
- 20 B. Johnson, S. Ho, Z. Xu, J. S. Yihan, S. Yip, E. Hingco and M. Leon: *J. Comp. Neurol.* **449** (2002) 180.
- 21 M. Imahashi and K. Hayashi: *Sens. Actuators, B* **166–167** (2012) 685.
- 22 M. Imahashi, K. Nakano and K. Hayashi: *Proc. Sensors, 2012 IEEE* (Taipei, Taiwan, 2012) p. 1080.
- 23 M. Imahashi and K. Hayashi: *J. Colloid Interface Sci.* **406** (2013) 186.
- 24 M. Imahashi, Y. Chiyoumaru and K. Hayashi: *Proc. Sensors, 2013 IEEE* (Baltimore, U.S.A., 2013) p. 258.