

A case study to standardize odor metadata obtained from coffee aroma based on E-nose using ISO/IEC 23005 (MPEG-V) for olfactory-enhanced multimedia

Jang-Sik Choi¹ and Hyung-Gi Byun^{2,*}

Abstract

Immersive multimedia comprising metadata for five senses can provide user experience by stimulating memory and sensation. In the case of olfactory-enhanced multimedia, a computer-generated smell is combined with additional media content to create a richer and/or more realistic experience for the user. Although several studies have been conducted on olfactory-enhanced multimedia using e-nose and olfactory display, their practical applications have been severely restricted owing to the absence of a related standard. This paper proposes a method to standardize odor metadata obtained from an e-nose system, which has been particularly used to acquire data from different coffee aromas. Subsequently, the data were transferred to an odor display applicable for olfactory-enhanced multimedia using the ISO/IEC 23005 (MPEG-V) data template.

Keywords : E-nose system, ISO/IEC 23005 (MPEG-V) data template, Olfactory-enhanced multimedia.

1. INTRODUCTION

Immersive multimedia is considered to be the next generation of multimedia services because it has significantly impacts the emotions, sense of presence, and engagement of a user [1]. The sense of presence and immersive experience in immersive multimedia are determined by sensory modalities and surround effect [2,3]. In particular, sensory modalities such as haptics, smell taste, and audio-visual content improve the quality of experience (QoE) of a user for immersive multimedia [4]. Therefore, it is important that a user perceives sufficient information through sensory modalities to experience immersive multimedia [5]. Among these sensory modalities, olfaction or the sense of smell, is an important perceptual function. Olfaction is known to enhance the quality of life and facilitate memory recall [6]. In addition, the olfactory stimuli enhance the productivity, alertness, and physical performance [7]. In the case of olfactory-enhanced

multimedia, a computer-generated smell is combined with additional media content to create a richer and/or more realistic experience for a user [8]. Several studies have been conducted to investigate the impact of olfaction on the user for olfaction-enhanced multimedia. Murray, Niall, et al. showed that the type of scent affects the QoE of a user and statistically significant differences between pleasant and unpleasant scents existed [9]. Ghinea, Gheorghita, and Oluwakemi Ademoye found that the association between scent and content has a significant impact on the user-perceived experience of olfactory-enhanced multimedia [10]. Additionally, they found that olfaction significantly improves the user multimedia experience [11]. Nakamoto et al. showed that presenting smells accurately enhanced the content [12]. The studies regarding olfaction-enhanced multimedia can be divided into two parts in terms of encoding and decoding of odor (scent) metadata (odor label, strength, concentration, and harmfulness). The first part is related to research using an olfactory display which can present odors to multiple users. This research focuses on adjusting important parameters which are associated with atomization, distribution, ventilation, and synchronization of odor [13-16]. The second part focuses on coding odor metadata obtained from an electronic nose. In these studies, an electronic nose (e-nose) was used to qualify and quantify an odor [17,18]. Most of the olfaction metadata coded from an e-nose was transferred to an olfactory display through a network and the olfactory display generated an odor corresponding to the odor recipe [19,20]. Several studies have been conducted to combine

¹ Department of Intelligent Crime Investigation, Donghae Police Station, 107, Cheongok-ro, Donghae-Si, Gangwon-Do, 25754, Korea

² Division of Electronics, Information & Communication Eng., Kangwon National University, 346, Joongang-ro, Samcheok-Si, Gangwon-Do, 25913, Korea

*Corresponding author: byun@kangwon.ac.kr

(Received: Jul. 10, 2021, Revised: Jul. 22, 2021, Accepted: Jul. 27, 2021)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

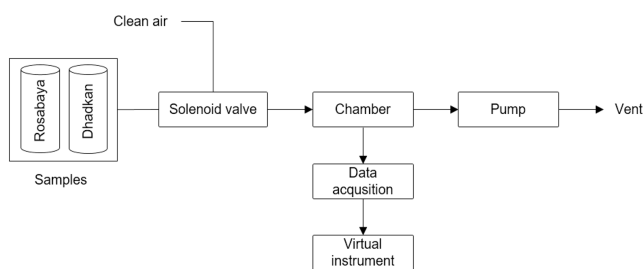


Fig. 1. Schematic diagram of E-Nose system.

olfaction with multimedia using an e-nose and olfactory display. However, their practical applications are restricted due to lack of device diversity, inability to utilize non-normative data templates, and absence of a generalized design for devices. Therefore, it is crucial to select a suitable standard that is universally accepted. ISO/IEC 23005 (MPEG-V), a representative standard related to immersive media, provides normative data templates and commands for the sensory effects of immersive multimedia [21-27]. In particular, this standard specifies information representation which can be used to control olfaction-related sensors (e-nose) and the actuators (olfactory display). This paper proposes a method to standardize odor metadata obtained from an electronic nose system which was particularly used to acquire data from different coffee aromas. Subsequently, the data were transferred to an odor display applicable for olfactory-enhanced multimedia using the MPEG-V (ISO/IEC 23005) data template. A virtual instrument and MPEG-V data templates used for e-nose system were presented in this paper.

2. METHODS

2.1 Electronic Nose (E-NOSE) System

An e-nose system consists of a solenoid valve, chamber, pump, data acquisition module, and a virtual instrument as shown in Fig. 1.

The solenoid valve and pump are used to flow clean air and target sample gases into a sensor array. The sensor array is comprised of four metal oxide semiconductor (MOS)-type sensors based on tin oxide (SnO₂). The sensor array was placed in a measurement chamber to improve the sensor response of signals in terms of stability, reproducibility and response time.

Clean air mixed with nitrogen (80%) and oxygen (20%) was used to stabilize the sensor array. Two bottles (500 mL) containing each sample were prepared. Table 1 shows the information of the samples: Dharkan and Rosabaya coffee capsules by Nespresso.

The experimental preparation included passing clean air

Table 1. Target samples.

| No. | Sample | Aromatic profile | Intensity |
|-----|----------|--|-----------|
| 1 | Dharkan | Long roasting/hints of bitter cocoa powder and toasted cereals | 11 |
| 2 | Rosabaya | Fruity-Wine | 6 |

through the chamber for 10 minutes before sampling. When the sensors were stabilized, the solenoid valve was opened to allow the scent of sample to flow inside the chamber. The sensor array responses to the samples were converted by a 10-bit ADC (analog-to-digital converter) of data acquisition module. The sampling time was 1 second. Each sensor response was transformed to a sensitivity value which is given by dR/R , where dR is the change in the resistance and R is the base resistance. The average value between two cursors as shown in Fig. 2 was extracted and used as a representative pattern for the input sample.

The representative pattern was normalized with respect to the overall sensor array response. Each normalized sensor response is indicative of a percentage of total response which is given by $(\frac{dR}{R})/\sum_{i=1}^n |dR/R|$, where n is the number of sensors. The normalized digital signals were transmitted through a universal asynchronous receiver-transmitter (UART) to the virtual instrument. The measurement was repeated five times for each sample. Ten patterns were extracted during the experiment. Eight patterns extracted were used as training data for coffee label recognition and the remaining patterns were used as test data.

The virtual instrument was developed to visually demonstrate the sensor array response of a sample, normalized pattern, recognition result of the normalized pattern, recognition result encoded to Enose Sensor Type, and the capability of e-nose encoded to Enose Sensor Capability Type. The eXtensible Markup Language (XML) was used for the encoding processing because the MPEG-V standard was originally written in XML Schema, which constitutes semi-structured data.

2.2 MPEG-V DATA TEMPLATES FOR E-NOSE SYSTEM

The odor metadata are encoded into Enose Sensor Type defined in MPEG-V part 5 [25]. A diagram of the data template is shown in Fig. 3.

Enose Sensor Type of MPEG-V consists of strength, harmfulness, chemical Gas Density, chemical Gas Density Unit, mono Chemical for a single gas, and mixture Chemical for a gas mixture. The strength describing odor intensity has one of the

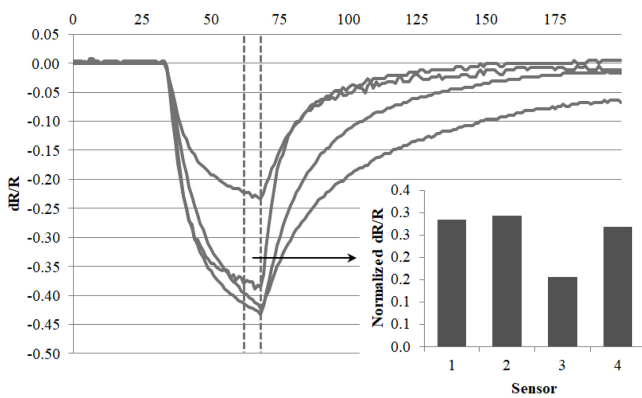


Fig. 2. Sensor array response and normalized pattern.

values listed in Table 2. Binary representation is used for reference software implementation defined in MPEG-V part 7 [27]. The harmfulness describing the harmful grade of the odor has one of the values listed in Table 3.

The chemical Gas Density is used to describe the concentration of the recognized single gas or mixture gas. The unit of chemical Gas Density is specified using a classification scheme (CS) term of Unit Type CS defined in MPEG-V part 6.

The recognized single gas and mixture gas are described using mono Chemical and mixtureChemical, respectively. Gas Type CS and Scent CS defined in MPEG-V part 6 provide CS terms for monoChemical and mixtureChemical. For describing the capability of E-Nose, MPEG-V standard provides a normative data template named EnoseCapabilityType [22] as shown in Fig. 4.

This data template allows a processor unit of an e-nose system to actuate sensors and actuators at an appropriate instant for data acquisition considering the operating condition of the e-nose system [28]. In addition, an auxiliary system cooperating with e-

Table 2. Terms for the strengths.

| No. | Term | Binary representation |
|-----|---------------|-----------------------|
| 1 | Imperceptible | 0000 |
| 2 | very_weak | 0001 |
| 3 | Weak | 0010 |
| 4 | Distinct | 0011 |
| 5 | Strong | 0100 |
| 6 | very_strong | 0101 |
| 7 | Intolerable | 0111 |

Table 3. Terms for the harmfulness.

| No. | Term | Binary representation |
|-----|-----------|-----------------------|
| 1 | no_hazard | 000 |
| 2 | Minimal | 001 |
| 3 | Moderate | 010 |
| 4 | Serious | 011 |
| 5 | Severe | 100 |

nose system can determine the capability of an e-nose through this template [29]. The Unit Type CS provides CS terms for the measurement Unit, and temp Unit. These elements are used to describe the unit of measurement and operating temperature of an e-nose, respectively. The warming-up time (in seconds) required to recognize scents is described as the warmUpTime.

The recognition time (in seconds) required to recognize scents is described as recognitionTime. The num of Recognition Scents is used to describe the number of recognizable scents that can be recognized through an e-nose. The max Operating Temp and min Operating Tem are used to describe the maximum temperature and minimum temperature (in Celsius), respectively, for the stable performance of an e-nose. The maximum humidity to ensure stable performance of E-Nose in terms of relative humidity (%) is described as max Operating Humid. The number of sensors present in an e-nose is described as num of Enose Sensors. The capability and recognizable scents for each gas sensor are described as Enose Sensors and recognition Scents, respectively.

3. RESULTS AND DISCUSSIONS

The main controller of the e-nose system for data acquisition was developed as shown in Fig. 5. The microcontroller unit (MCU) was used to control the sensors and actuators. The four MOS-type sensors were connected to each ADC channel of the main controller. The pump was connected to OC1A (PB5) of MCU and controlled by PWM (pulse-width modulation) method.

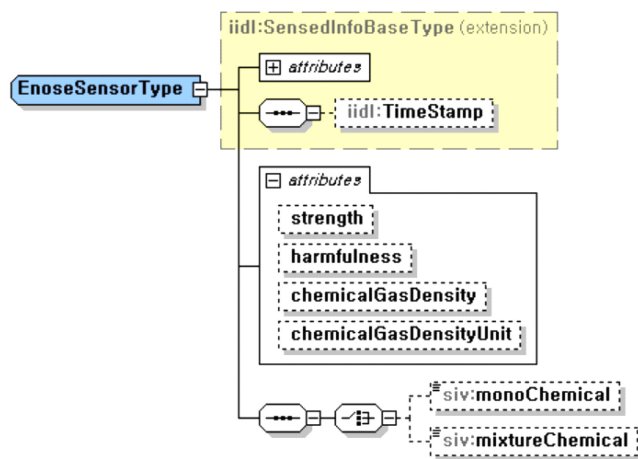


Fig. 3. Diagram of Enose Sensor Type.

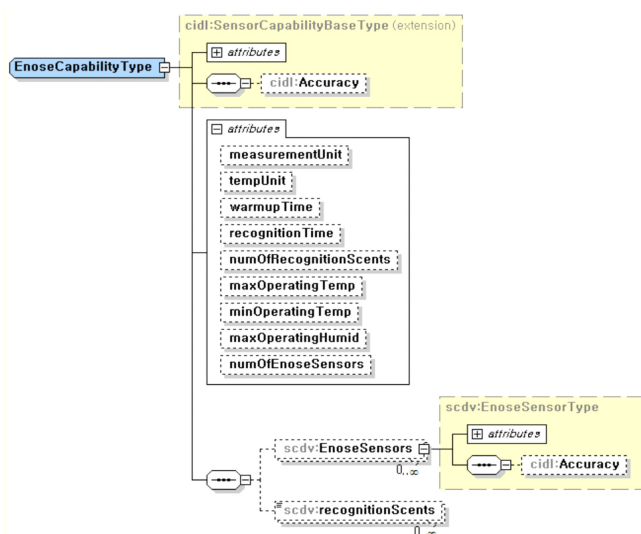


Fig. 4. Diagram of Enose Capability Type.

The solenoid valve was connected to the MCU through a Darlington transistor. A USB to UART cable was used to transfer the response of the sensor array to the virtual instrument.

The eight normalized patterns obtained from the e-nose system or the training data, are listed in Table 4. The traditional backpropagation (BP) algorithm was used for coffee label recognition (0: Dharkan, 1: Rosabaya). The network results for the training data are listed in Table 5. The trained network accurately recognized the coffee labels of all samples. The trained networks were embedded in the virtual instrument to recognize the coffee label of the test data.

Fig. 6 shows the implemented virtual instrument graphical interface. The response of sensor array to the sample is displayed at the left-top side. The normalized pattern for the response is displayed at the left-bottom side. The recognition result, strength,

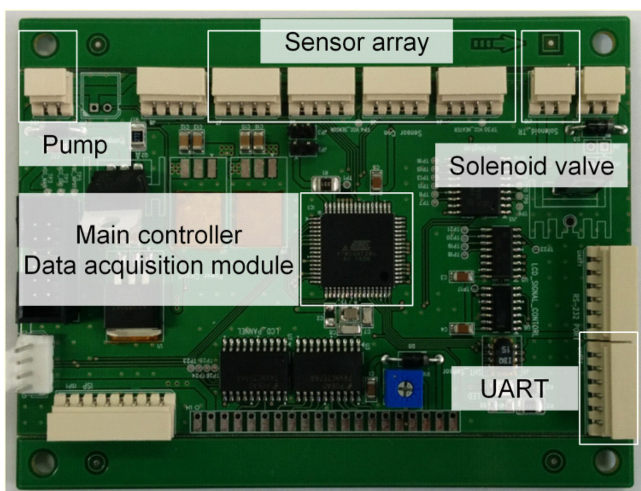


Fig. 5. Main controller board of E-Nose system.

Table 4. Normalized Patterns for Training Data.

| No. | Sensor 1 | Sensor 2 | Sensor 3 | Sensor 4 | Sample |
|-----|----------|----------|----------|----------|----------|
| 1 | 0.2804 | 0.2928 | 0.1581 | 0.2687 | Dharkan |
| 2 | 0.2821 | 0.2957 | 0.1580 | 0.2642 | Dharkan |
| 3 | 0.2839 | 0.2950 | 0.1567 | 0.2643 | Dharkan |
| 4 | 0.2835 | 0.2941 | 0.1605 | 0.2620 | Dharkan |
| 5 | 0.2786 | 0.2757 | 0.1466 | 0.2990 | Rosabaya |
| 6 | 0.2783 | 0.2761 | 0.1458 | 0.2998 | Rosabaya |
| 7 | 0.2783 | 0.2746 | 0.1461 | 0.3009 | Rosabaya |
| 8 | 0.2794 | 0.2742 | 0.1470 | 0.2994 | Rosabaya |

Table 5. Recognition Results for Training Data.

| No. | Coffee Aroma | Target | Recognition result |
|-----|--------------|--------|--------------------|
| 1 | Dharkan | 0 | 0.0241 |
| 2 | Dharkan | 0 | 0.0077 |
| 3 | Dharkan | 0 | 0.0087 |
| 4 | Dharkan | 0 | 0.0052 |
| 5 | Rosabaya | 1 | 0.9826 |
| 6 | Rosabaya | 1 | 0.9851 |
| 7 | Rosabaya | 1 | 0.9887 |
| 8 | Rosabaya | 1 | 0.9857 |

chemical gas density (unit), and harmfulness are displayed at the right-top side. At the right-bottom side, the two tree views display information about the recognized odor metadata and the capability of the e-nose, which are encoded in each template of MPEG-V.

Fig. 7-8 and Table 6 show and list results of two coffee labels, which are Dharkan and Rosabaya, classification and encoded data template for the test set. The strength and harmfulness were determined by a look-up table containing their values according to the recognized coffee label. In particular, the strength values inside the look-up table were defined considering the values of intensity listed in Table 1. Since experimental results were not available to estimate the concentration, the corresponding values for Chemical Gas Density and Chemical Gas Density Unit were treated as not applicable (N/A). The odor metadata for the Dharkan test sample and capability of the e-nose system were encoded to Enose Sensor Type and Enose Sensor Capability Type and then displayed hierarchically on the tree viewers. The odor metadata comprised of the recognized coffee label (Dharkan), strength (very_strong), harmfulness (no_hazard) was listed on the corresponding tree viewer.

The representative patterns for the test samples and their recognition results are listed in Table 6. The trained network accurately classified the coffee labels of the test set.

Although sensor 4 exhibits a relatively good selectivity to

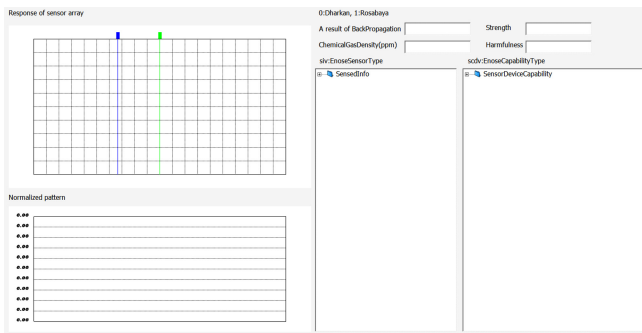


Fig. 6. VI (Virtual Instrument) graphical interface.

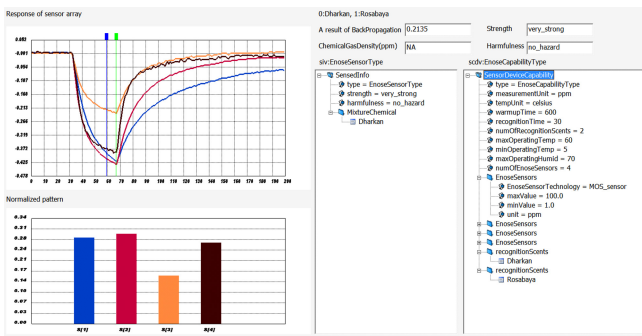


Fig. 7. VI and results for Dharkan test sample.

Table 6. Test Data and Recognition Results.

| No. | Sensor 1 | Sensor 2 | Sensor 3 | Sensor 4 | Sample | Recognition result |
|-----|----------|----------|----------|----------|----------|--------------------|
| 1 | 0.2841 | 0.2923 | 0.1562 | 0.2674 | Dharkan | 0.2135 |
| 2 | 0.2785 | 0.2744 | 0.1477 | 0.2994 | Rosabaya | 0.9857 |

Rosabaya, there are no significant differences between the response of sensors for the samples, Dharkan, and Rosabaya. Therefore, an investigation on various gas sensors to improve selectivity is additionally required.

E-Nose has been widely used to assess the quality of coffee [30]. In particular, artificial neural network (ANN) is generally used to distinguish the coffee label. In this study, the BP algorithm, which is an ANN algorithm, was used. Although E-Nose has applications in different fields, the diffusion of e-nose in real-life applications is still limited due to absence of standards related to data template, interface, calibration techniques, and reference materials. Under such circumstances, European e-nose standardization initiative (Nose II) proposed standard data formats and utilization of IEEE 1451, a set of smart transducer interface standards using open, common, network-independent communication interfaces [31]. In addition, a systematic way to develop a standard e-nose was proposed [32]. In this article,

MPEG-V standard data templates associated with e-nose were provided. In addition, the e-nose system, which adopted the data templates, was presented including the virtual instrument which can recognize the coffee label and encode odor metadata to the data templates.

4. CONCLUSIONS

In this paper, MPEG-V standard data templates related to e-nose for olfactory-enhanced media, e-nose system adopted data templates, and virtual instrument used for the recognition of coffee label and encode odor metadata to the data templates, were presented.

The olfactory-enhanced media requires standard data templates of odor metadata, which can be easily combined with existing visual-audio content rather than data templates describing the detailed information (e.g., output voltage level of each gas sensor) detected through an e-nose. The detailed information would be interpreted differently in recognition processes. Therefore, it is important to provide a generic standard metadata for future content-based multimedia application. The research for the standardization of e-nose is crucial because a normative model, interface, and data format can accelerate the commercialization and ensure interoperability of the devices, and enhance the user experience.

REFERENCES

- [1] N. Murray, B. Lee, Y. Qiao, and G. Muntean, "Olfaction-enhanced multimedia: A survey of application domains, displays, and research challenges", *ACM Computing Surveys (CSUR)*, Vol. 48, No. 2, pp. 1-34, 2016.
- [2] T. Chambel, "Interactive and Immersive Media Experiences", *Proc. of WebMedia'16*, pp. 1, Teresina, PI, Brazil, 2016
- [3] P. Viana, T. Chambel, V. Michael Bove Jr., S. Strover, and G. Thomas, "Guest Editorial: Immersive Media Experiences", *Multimedia Tools Appl.*, Vol. 75, pp. 12285-12290, 2016.
- [4] Z. Yuan, S. Chen, G. Ghinea, and G. Muntean, "User quality of experience of mulsemmedia applications", *ACM Transactions on Multimedia Comput. Commun. Appl.*, Vol. 11, No. 1s, pp. 1-19, 2014.
- [5] N. Ransinghe, K. Lee, and E. Do, "The sensation of taste in the future of immersive media", *Proc. of ImmersiveMe'14*, pp. 7-11, Orlando, FL, USA, 2014.
- [6] M. Howell, N. Herrera, A. Moore, and R. McMahan, "A reproducible olfactory display for exploring olfaction in

- immersive media experiences”, *Multimedia Tools and Applications*, Vol. 75, pp. 12311-12330, 2016.
- [7] D. Washburn, L. Jones, R. Satya, and C. Bowers, “Olfactory use in virtual environment training”, *Modeling & Simulation*, Vol. 2, No. 3, pp. 19-25, 2003.
- [8] O. Ademoye, and G. Ghinea. “Information recall task impact in olfaction-enhanced multimedia”, *ACM Transactions on Multimedia Computing, Communications, and Applications*, Vol. 9, No. 3, pp. 1-16, 2013.
- [9] N. Murray, B. Lee, Y. Qiao, and G. Miro-Muntean, “The impact of scent type on olfaction-enhanced multimedia quality of experience.”, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 47, No. 9, pp. 2503-2515, 2016.
- [10] G. Ghinea and O. Ademoye. “User perception of media content association in olfaction-enhanced multimedia”, *ACM Transactions on Multimedia Computing, Communications, and Applications*, Vol. 8, No. 4, pp. 1-19, 2012.
- [11] G. Ghinea and O. Ademoye. “The sweet smell of success: Enhancing multimedia applications with olfaction”, *ACM Transactions on Multimedia Computing, Communications, and Applications*, Vol. 8, No. 1, pp. 1-17, 2012.
- [12] T. Nakamoto, S. Otaguro, and M. Kinoshita, “Cooking up an interactive olfactory game display”, *IEEE Computer Graphics and Applications*, Vol. 28, No. 1, pp. 75-78, 2008.
- [13] K. Hashimoto and T. Nakamoto, “Stabilization of SAW atomizer for a wearable olfactory display”, *Proc. of IEEE International Ultrasonic Symposium*, pp. 1-4, Taipei, Taiwan, 2015.
- [14] T. Nakamoto, K. Hashimoto, T. Aizawa, and Y. Ariyakul, “Multi-compont olfactory display with a SAW atomizer and micropumps controlled by a tablet PC”, *Proc. of IEEE International Frequency Control Symposium*, pp. 1-4, Taipei, Taiwan, 2014.
- [15] H. Matsukura, T. Yoneda, and H. Ishida, “Smelling Screen: Development and Evaluation of an Olfactory Display System for Presenting a Virtual Odor Source”, *IEEE Trans. On Visualization and Computer Graphics*, Vol. 19, No. 4, pp. 606-615, 2013.
- [16] A. Fukasawa and K. Okada, “Olfactory Measurement Method at Health checkup with Olfactory Display using Pulse Ejection”, *International Journal of Information Society*, Vol. 5, No. 1, pp.13-19, 2013.
- [17] P. Somboon, B. Wyszynski, and T. Nakamoto. “Novel odor recorder for extending range of recordable odor”, *Sensors and Actuators B: Chemical*, Vol. 121, No. 2, pp. 583-589, 2007.
- [18] M. Son, J. Lee, H. Ko, and T. Park, “Bioelectronic Nose: An Emerging Tool for Odor Standardization”, *Trends in Biotechnology*, Vol. 35, No. 4, pp. 301-307, 2017.
- [19] T. Nakamoto, Eds., *Essentials of machine olfaction and taste*. John Wiley & Sons, Singapore, pp. 247-314, 2016.
- [20] K. Toko, Eds., *Biochemical sensors: mimicking gustatory and olfactory senses*. Pan Stanford, pp. 285-304, 2013.
- [21] ISO/IEC 23005-1:2016, Information technology - Media context and control - Part 1: Architecture, July 2016.
- [22] ISO/IEC 23005-2:2016, Information technology - Media context and control - Part 2: Control information, Mar. 2016.
- [23] ISO/IEC 23005-3:2016, Information technology - Media context and control - Part 3: Sensory information, July 2016.
- [24] ISO/IEC 23005-4:2016, Information technology - Media context and control - Part 4: Virtual world object characteristics, Mar. 2016.
- [25] ISO/IEC 23005-5:2016, Information technology - Media context and control - Part 5: Data formats for interaction devices, Mar. 2016.
- [26] ISO/IEC 23005-6:2016, Information technology - Media context and control - Part 6: Common types and tools, Mar. 2016.
- [27] ISO/IEC 23005-7:2014, Information technology - Media context and control - Part 7: Conformance and reference software, Jan. 2014.
- [28] J. Choi, S. Chang, H. Lee, and H. Byun, “Olfactory Interaction based on ISO/IEC 23005 Standard.”, *J. Sens. Sci. Technol.* Vol. 26, No. 5, pp. 297-300, 2017.
- [29] J. Choi, J. Jeon, S. Chang, H. Lee, and H. Byun, “A normative template describing capability of E-Nose in MPEG-V”, *Proc. of 7th GOSPEL*, Seoul, Korea, Nov. 2017
- [30] J. Gardner, W. Shurmer, and T. Tan. “Application of an electronic nose to the discrimination of coffees”, *Sensors and Actuators B: Chemical*, Vol. 6, No. 1-3, pp. 71-75, 1992.
- [31] N. Ulivieri, C. Distanto, T. Luca, S. Rocchi, and P. Siciliano, “IEEE1451. 4: A way to standardize gas sensor”, *Sensors and Actuators B-Chemical*, Vol. 114, No. 1, pp. 141-151, 2006.
- [32] N. Kaja, L. Alazzawi, and G. Raishouni, “E-Nose, Developing a Standard Product”, *Proc. of ISEANS*, pp. 102-109, Macau, China, 2013.