# International Journal of Informatics Society

16/06 Vol. 8 No.1 ISSN 1883-4566

Informatics Society

**Aims and Scope**

The purpose of this journal is to provide an open forum to publish high quality research papers in the areas of informatics and related fields to promote the exchange of research ideas, experiences and results.

Informatics is the systematic study of Information and the application of research methods to study Information systems and services. It deals primarily with human aspects of information, such as its qu ality and value as a resource. Informatics also referred to as Information science, studies t he structure, algorithms, behavior, and interactions of natural and a rtificial systems that store, process, access and communicate information. It also develops its own conceptual and theoretical foundations and utilizes foundations developed in other fields. The advent of computers, its ubiquity and ease to use has led to th e study of info rmatics that has computational, cognitive and social aspects, including study of the social impact of information technologies.

The characteristic of informatics' context is amalgamation of technologies. For creating an informatics product, it is necessary to integrate many technologies, such as mathematics, linguistics, engineering and other emerging new fields.

# Guest Editor’s Message

Hiroshi Ishikawa
Guest Editor of Twenty-second Issue of International Journal of Informatics Society

We are delighted to have the twenty-second issue of the International Journal of Informatics Society (IJIS) published. This issue includes selected papers from the Ninth International Workshop on Informatics (IWIN2015), which was held at Amsterdam, Netherlands, Sep. 6-9, 2015. The workshop was the ninth event for the Informatics Society, and was intended to bring together researchers and practitioners to share and exchange their experiences, discuss challenges and present original ideas in all aspects of informatics and computer networks. In the workshop 24 papers were presented in five technical sessions. The workshop was successfully finished with precious experiences provided to the participants. It highlighted the latest research results in the area of informatics and its applications that include networking, mobile ubiquitous systems, data analytics, business systems, education systems, design methodology, intelligent systems, groupware and social systems.

Each paper submitted IWIN2015 was reviewed in terms of technical content, scientific rigor, novelty, originality and quality of presentation by at least two reviewers. Through those reviews 14 papers were selected for publication candidates of IJIS Journal, and they were further reviewed as a Journal paper. This volume includes five papers among the accepted papers, which have been improved through the workshop discussion and the reviewers’ comments.

We publish the journal in print as well as in an electronic form over the Internet. We hope that the issue would be of interest to many researchers as well as engineers and practitioners over the world.

**Hiroshi Ishikawa** received the B.S. and Ph.D degrees in Information Science from the University of Tokyo. After working for Fujitsu Laboratories and being a full professor of Shizuoka University, he is now a full professor of Tokyo Metropolitan University from April, 2013. His research interests include database, data mining, social media, and big data. He has published actively in international, refereed journals and conferences, such as ACM TODS, IEEE TKDE, VLDB, IEEE ICDE, and ACM SIGSPATIAL. He has authored some books, which include books entitled *Social Big Data Mining* (CRC), *Object-Oriented Database System* (Springer-Verlag), *databases* (Morikita Publishing), *an introduction to social big data science* (Corona Publishing), and *data mining and collective intelligence* (Kyoritsu Shuppan). He received the Sakai Memorial Distinguished Award from IPSJ (Information Processing Society of Japan) and the Director General Award from Science and Technology Agency of Japan. He was an invited professor at the Polytechnic School of the University of Nantes, France, twice. He was a trustee board member of the Database Society of Japan, an editorial board member of VLDB Journal, the chairman of the SIG on Database Systems of IPSJ, and an editor-in-chief of IPSJ Trans. on Databases. He is fellows of IPSJ and IEICE (The Institute of Electronics, Information and Communication Engineers) and members of ACM and IEEE.

# Improvement of Accuracy based on Multi-Sample and Multi-Sensor in the Gait-based Authentication using Trouser Front-Pocket Sensors

Shinsuke Konno*, Yoshitaka Nakamura**, Yoh Shiraishi**, and Osamu Takahashi**

* National Institute of Technology, Hakodate Collage, Japan
** School of Systems Information Science, Future University Hakodate, Japan
skonno@hakodate-ct.ac.jp
{y-nakamr, siraisi, osamu}@fun.ac.jp

***Abstract*** - Nowadays, smartphones are equipped with various personal authentication functions to prevent imposters from misusing. However, low utilization rate of personal authentication functions caused by the inconvenient operations becomes the problem. To reduce the inconvenience, authentication methods based on behavior characteristics have been studied. Gait-based authentication is one of them. This authentication method automatically identifies individuals based on walking motions measured by smartphone sensors. The purpose of this study is to find the appropriate methods in each authentication process and to improve the authentication accuracy. For these purpose, we tested the various distance calculation methods, and four fusion algorithms. As this result, we found the appropriate methods in each process. Furthermore, by applying these appropriate methods to the authentication system, the authentication accuracy was significantly improved in comparison with previous studies.

***Keywords***: gait-based authentication, acceleration sensor, gyro sensor, dynamic time warping, score level fusion

## 1 INTRODUCTION

The use of portable terminals such as smartphones has increased in recent years and can be expected to increase in the future. Though initially used mainly for telephone calls, text messages, and browsing, nowadays such devices also serve many other important functions, such as making payments and storing private and business information. Accordingly, smartphones and other portable terminals have become loaded with various functions intended to prevent misuse by imposters. One of these is personal authentication, for which there are many methods, such as personal identification numbers (PINs) and passwords. Recently, methods such as pattern locks, which are more difficult for an imposter to break, has been incorporated into the devices. Furthermore, terminals with biometric identification using fingerprints or face images have been developed. As these examples illustrate, developers are making serious efforts to prevent improper use of these devices. On the other hand, there are reports and news items showing that approximately 50% of users do not use personal authentication on their devices, perhaps because they feel that the authentication methods are too difficult to use. Previous studies proposed easier authentication methods by extracting individual features of device operation, such as swinging the terminal or tapping on the display. However these methods require conscious action, so they cannot perform authentication in the background.

On the other hand, it is conceivable that individual authentication might be established through daily repeated activities. With such a method, users can unlock their terminal without conscious authentication operations. Gait-based authentication is one example of an unconscious method of this type. The Ministry of Health, Labour and Welfare reported that Japanese adult men walk an average of 7,099 steps in a day. Furthermore, we think that walking is performed in various situations. If gait-based authentication was established, the inconvenience users feel in individual authentication would be reduced. There are many scenes when we want to unlock our smartphones quickly and automatically in daily life. Examples are automatic-ticket gate and payment with IC card of smartphone at the cash desk. Furthermore, this authentication system can identify the owner without conscious operation repeatedly. Therefore, this authentication system can use as the theft detecting system. This authentication system can develop with the simple sensors. So, there is a research which proposed that this authentication system was built into smart key system for cars.

We work with multi-modal authentication to improve authentication performance by combining multiple methods in individual authentication [1].

Fernand et al. [2] combined faces and fingerprints to improve accuracy. Zhou et al. [3] combined features of side face and gait using principal component analysis to identify people, and many other researchers have also attempted to improve accuracy using biometric authentication.

However, wearing multiple sensors on various body parts sacrifices convenience, the advantage of gait authentication. For this reason, we adopt a method that combines multiple sensor methods measuring the same body parts using multiple sensors, and a multi-sample method that measures a modality several times to improve performance. It is possible to equip a terminal with multiple sensors, enabling us to authenticate using multiple sensors without imposing a burden on users.

In this study, we use two sensors (a three-axis acceleration sensor and a three-axis gyro sensor) to measure human walking motion. We show that the proposed method, which combines distance information recorded by these two sensors, improves authentication accuracy in comparison with previous studies.

Table 1: Summary of gait-based authentication work.

| work | position | Sensor |
|---|---|---|
| Mäntyjärviet et al. [4] | belt | acceleration |
| Gafurov et al. [5] | hip | acceleration |
| Gafurov et al. [6] | ankle | acceleration |
| Gafurov et al. [7] | trouser pocket | acceleration |
| Gracian et al. [8] | belt | acceleration |
| Derawi et al. [9] | belt | acceleration |
| Soumik et al. [10] | eight-joints | rotation angle |

# 2 RELATED WORK

There has been related work on authentication using wearable sensors for walking motions. The most common approach is to use acceleration sensors fixed at the user's waist. We explain the details of that work in this section.

## 2.1 Position of Sensors

Table 1 summarizes the related work. These studies explored features and authentication methods primarily to improve performance. However, they did not investigate which sensor positions would be acceptable for daily use. Those studies measured mainly using devices attached on the belt on the middle or side of the waist, and authenticated subjects based on the acceleration signals. This requires using a smartphone case such as a holster for attaching the terminal to the waist. Users might find this unacceptable, because gait authentication then requires them to have the container with them. Consequently, we decided that the trouser front-pocket might be acceptable to users, because they can then have the terminal without using special tools, and we investigated performance improvement in this position. The study in [7] examined this position. This study aims to improve authentication performance in comparison to that previous study.

## 2.2 Fusion of Multiple Sensors

Many acceleration-based approaches to gait-based authentication have been explored.

Mäntyjärviet et al. [4] proposed three authentication methods: fast Fourier transform, correlation, and statistical features.

Gafurov et al. [5]-[7] studied methods based on acceleration, and made measurements by using acceleration sensors on various parts of subject's bodies. They used a template signal and multiple time-normalized signals, with the acceleration sensor placed in the trouser front pocket [7].

Gracian et al. [8] devised the feature of gait acceleration for user authentication.

Derawi et al. [9] proposed a multi-sampling method that authenticated using multiple signals from both templates and inputs. Their method calculated distances of all combinations of templates and inputs with dynamic time warping (DTW).

Soumik et al. [10] measured walking motions with eight angle sensors attached on various joints.

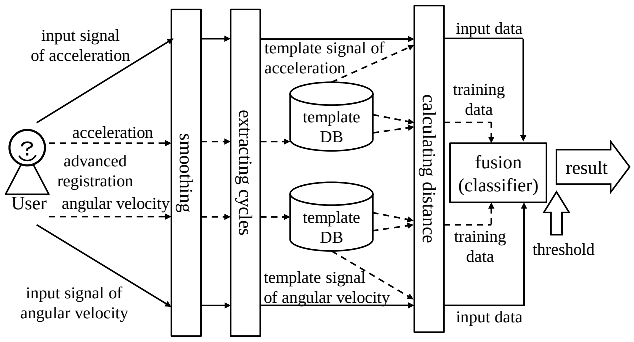



Figure 1: Overview of the proposed system.

To the best of our knowledge, there are no studies on the fusion of multiple sensors placed in a trouser front pocket. In this study, to improve authentication accuracy, we propose a method of fused distances based on acceleration and angular velocity placed in a trouser front pocket.

# 3 PROPOSED METHOD

## 3.1 Overview

Figure 1 shows an overview of the proposed authentication method. This method authenticates users based on distances among input and template signals in the time domain. Template signals are registered in the database in advance.

Gait signals show similar waveforms repeatedly for each subject. We expect that when genuine, the distance between input and template signals is small, and when from an imposter, the distance is greater. Consequently, we considered that it is possible to generate a common classifier among all users for authentication. When we fuse distances calculated from multiple sensors, we use a common classifier in all subjects.

When the system employed multi-sample method, users need to walk long distances for multiple input signals to be used for authentication. Therefore, we consider that multi-sample authentication must calculate the distance between an input signal and the multiple template signals registered.

## 3.2 Gait Recognition and Quasi-Periodic Signal Extraction

We attached a sensor unit whose x-, y-, and z-axis detected vertical, sideway, and forward-backward acceleration, respectively, in standing posture. The direction of each axis is shown in Fig. 2. Each signal of a gyro sensor detects an angular velocity whose rotary axis is identical to each axis of an acceleration sensor. Each subject wore a sensor unit attached to a belt with hook and loop fastener. This unit was placed on the front of the left femur area.

Examples of three-axis acceleration and three-axis angular velocity are shown in Fig. 3 and 4. During walking, the acceleration and gyro sensors measured similar waveforms repeatedly. These signals are quasi-periodic signals with no equalization of cycles and amplitudes. The length of a gait cycle is two steps. The gait cycle consists of four periods, double limb support (in left stance phase), single limb support (in right swing phase), double limb support (in right

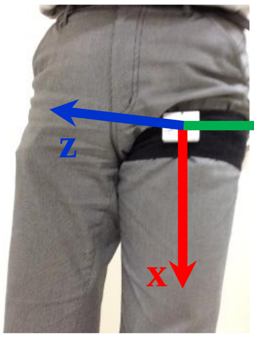



Figure 2: Directions of three axes.

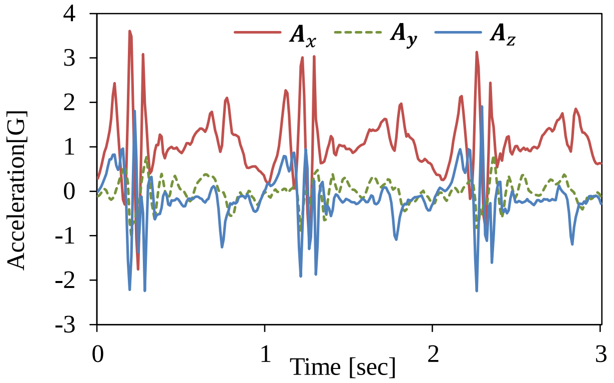



Figure 3: Example of gait signals from three-axis acceleration sensor.

stance phase) and single limb support (in left swing phase). We walk forward by repeating the four periods. If we change the start period of the signal extraction for each user to obtain the input signals and the template signals whose lengths are two steps, we extract the signals with different order of periods for each user. As a result, we obtain their signals with different waveform among users and may achieve good performance seemingly in authentication. To prevent influence on authentication accuracy by different waveform for each user, we decided to extract their quasi-periodic signals with the same order of the gait periods to all users. For this reason, we conducted a preliminary experiment to investigate the relation between walking motion and six-axis signals. To measure the time from a left heel touching ground to its rising from the ground, two force sensors synchronized with the sensor unit were attached to their left toe and heel. Examples of the acceleration along the x-axis and the signal of the force sensor are shown in Fig. 5. The graph shows that the time when the acceleration becomes a local maximum is approximately equal to the time when the value of the force sensor under the left heel begins to increase. This result indicates that the time when the acceleration becomes a local maximum coincides with the time when the left heel lands on the ground. Hence, we extract these quasi-periodic signals of the same walking motion period in the different subjects using the following method.

### 3.2.1 Walking Detection

In this study, we use a threshold in vertical acceleration to detect walking start time based on previous research [7].

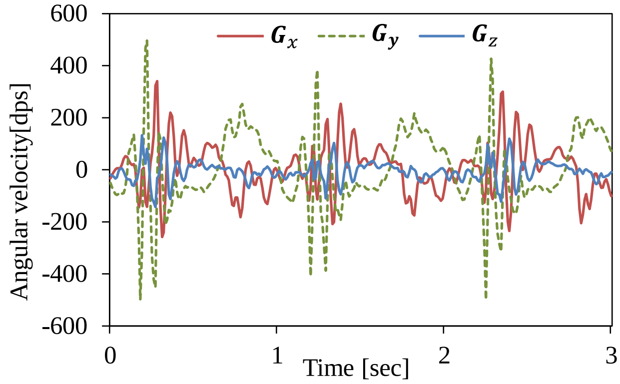



Figure 4: Example of gait signals from three-axis angular gyro sensor.

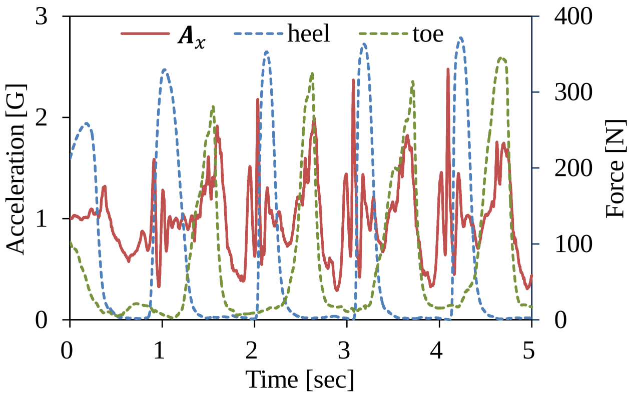



Figure 5: Example of the vertical acceleration signal and force signal.

Before beginning, all signals were smoothed using a Savitzky–Golay filter [11]. We look for the time $t_s$ when the acceleration is greater than 1.2 G from the start of this quasi-periodic signal extraction method.

### 3.2.2 Quasi-Periodic Signal Extraction

After detection, we extract quasi-periodic signals measuring the period between left-heel landings on the ground. The extraction process with x-axis acceleration $\boldsymbol{A}_x$ is as follows:

1) We search for the maximum time $T_0$ within two seconds after walking detection $t_s$. We selected $T_0$ as the start time of cycle $\boldsymbol{C}_0$. For reference, the process of detecting $T_0$ is shown in Fig. 6.
2) To find the end time of $\boldsymbol{C}_0$, we search for all times of local maxima $\boldsymbol{t}_1 = \{t_{11}, t_{12}, t_{13} \cdots\}$ from 0.7 to 1.3 s after $T_0$ from $\boldsymbol{A}_x$.
3) We extract subsets $\boldsymbol{s}_0$ that are 0.6 s of the signal. $T_0$ is the middle time of subset $\boldsymbol{s}_0$. In the same way, each element in $\boldsymbol{t}_1$ is the middle time of subsets $\boldsymbol{S}_1 = \{\boldsymbol{s}_{11}, \boldsymbol{s}_{12}, \boldsymbol{s}_{13} \cdots\}$, whose subsets are extracted as 0.6 s signals. For example, $\boldsymbol{s}_{11}$ are 0.6s of extracted signal whose middle time is $t_{11}$. We calculate values $\boldsymbol{N}_1 = \{N_{11}, N_{12}, N_{13} \cdots\}$ of the normalized cross correlation (NCC) among $\boldsymbol{s}_0$ and each element in $\boldsymbol{S}_1$. We take the middle time of $\max(\boldsymbol{N}_1)$ as the start time $T_1$ of the next cycle $\boldsymbol{C}_1$. Figure 7 shows an example of detecting $T_1$. As a result of the calculations using NCC, we take $t_{13}$ as $T_1$.

Through this processing, we obtain one cycle $\boldsymbol{C}_0$ by extracting $\boldsymbol{A}_x$ from $T_0$ to $T_1$.

4) Next, we search for all times of local maxima $\boldsymbol{t}_2 = \{t_{21}, t_{22}, t_{23} \cdots\}$ from 0.7 to 1.3 s after $T_1$. We extract subsets of signal $\boldsymbol{S}_2 = \{\boldsymbol{s}_{21}, \boldsymbol{s}_{22}, \boldsymbol{s}_{23} \cdots\}$ from $T_1$ to each $\boldsymbol{t}_2$. We calculate distances $\boldsymbol{d}_2 = \{d_{21}, d_{22}, d_{23} \cdots\}$ between $\boldsymbol{C}_0$ and each $\boldsymbol{S}_2$ using DTW. To eliminate the effect of differences in signal length, we divided each distance by the total length of $\boldsymbol{C}_0$ and each $\boldsymbol{S}_2$. We take the time $t_{22}$ of minimum distance as the start time $T_2$ of the next cycle $\boldsymbol{C}_2$.
5) After the time $T_n$ of minimum distance is calculated using DTW among $\boldsymbol{C}_{n-1}$ and $\boldsymbol{S}_n$, we begin searching for the next start time $T_{n+1}$ by repeating step 4) .
6) When forward searching is completed, we repeat the process by searching backward at $T_0$.
7) The end time of an extracted cycle duplicates the start time of the next cycle. Hence, we eliminated amplitude of end time from each extracted cycle.
8) When we observed the extracted signals, we found that those near the signals of starting to walk had a large distortion as compared with other signals. Based on the result of analysis, the variance of each signal with a large distortion is smaller than the variance of other signals. Hence, we searched for the first distorted signals whose variance was greater than the threshold 0.09. We assumed that the signals used for authentication were signals subsequent to it. Examples of the variance from extracted signals are shown in Fig. 9. In this x-axis acceleration, we took the signals to be used for authentication as the cycles after $C_0$ . We recorded the starting times of extracted cycles, and extracted signals for the other two-axis acceleration and three-axis angular velocity using the same starting time.

The x-axis signals extracted by this method are shown in Fig. 10 and 11. Figure 10 shows two quasi-periodic signals extracted from the same subject. In Fig. 11, the solid and dotted lines indicate extracted quasi-periodic signals from different subjects.

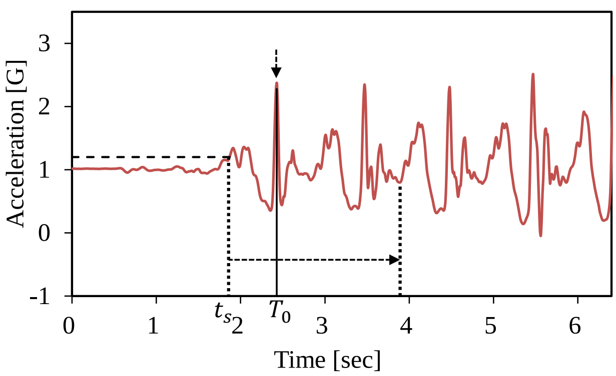



Figure 6: Example of detection of start time $t_0$ of cycle $\boldsymbol{C}_0$.

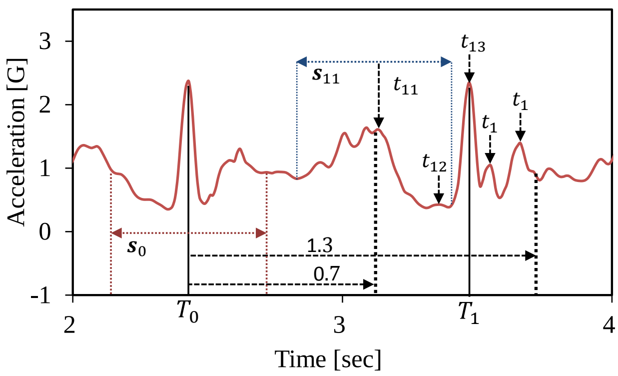



Figure 7: Example of detection of start time $t_1$ of cycle $\boldsymbol{C}_1$. Lower-arrows indicate the local maxima between 0.7 and 1.3 s after $T_0$.

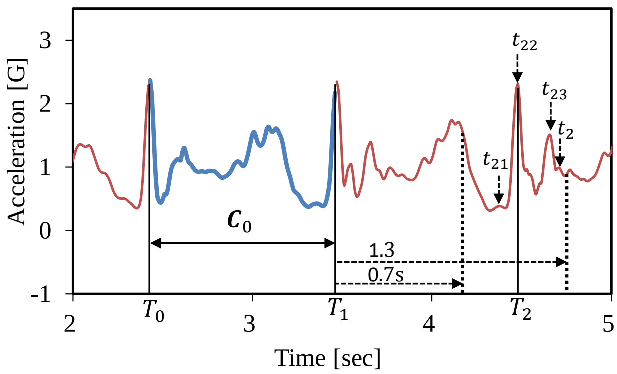



Figure 8: Example of extracted cycle $\boldsymbol{C}_0$ and local maximum $\boldsymbol{t}_2$.

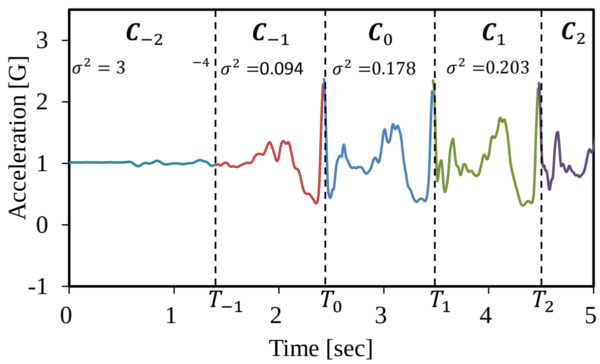



Figure 9: Example of extracted cycles and their variances.

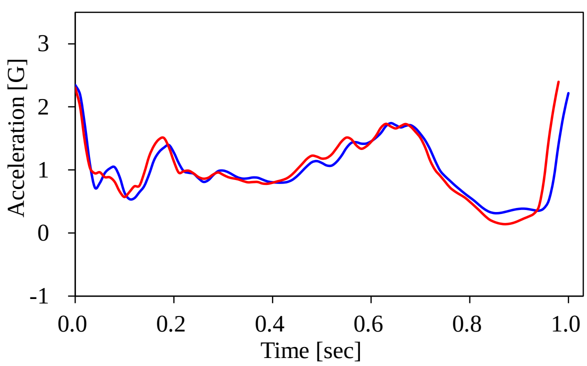



Figure 10: Extracted signals from same subject.

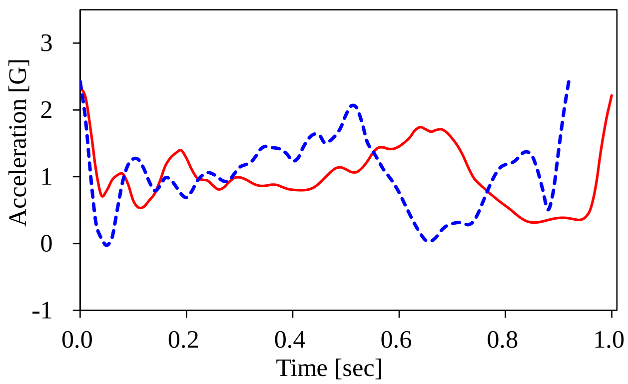



Figure 11: Extracted signals from two subjects.

## 3.3 Distance Calculation Methods

To find the better method, we tested DTW and linear interpolation which equalize the length of two signals. These are frequently used for calculating dissimilarity as a distance between time series data. Let $\boldsymbol{X} = \{x(i) | i = 1,2,\cdots,m\}$, $\boldsymbol{Y} = \{y(j) | j = 1,2,\cdots,n\}$ be time series data. The distance by DTW between $\boldsymbol{X}$ and $\boldsymbol{Y}$ is defined as

$$D_{DTW}(\boldsymbol{X},\boldsymbol{Y}) = f(m,n)$$
$$f(i,j) = \min \begin{cases} f(i-1,j-1) + dist(x(i),y(j)) \\ f(i,j-1) + dist(x(i),y(j)) + GP \\ f(i-1,j) + dist(x(i),y(j)) + GP \end{cases}$$
$$f(0,0) = 0, \quad f(1,0) = f(0,1) = \infty$$

where $D_{DTW}(\boldsymbol{X},\boldsymbol{Y})$ is the distance calculated using DTW, $m$ and $n$ are the number of lengths in signals $\boldsymbol{X}$ and $\boldsymbol{Y}$, and $GP$ is a gap penalty in the case of non-linear extension. We adopted the different distance calculation method for each sensor. The distance calculation function is substituted into $dist(x(i),y(j))$. Next, to adapt the differences of signal length by differences of walking speed, we normalized the distance by dividing the total length of the two signals. The normalized distance $D(\boldsymbol{X},\boldsymbol{Y})$ is calculated as

$$D(\boldsymbol{X},\boldsymbol{Y}) = \frac{D_{DTW}(\boldsymbol{X},\boldsymbol{Y})}{m+n}$$

Similarly, in the case of linear interpolation, the length of $\bar{\boldsymbol{X}}$ and $\bar{\boldsymbol{Y}}$ was equalized to the longer length between $\boldsymbol{X}$ and $\boldsymbol{Y}$ using linear interpolation to $\boldsymbol{X}$ and $\boldsymbol{Y}$. The normalized distance $D(\boldsymbol{X},\boldsymbol{Y})$ is calculated as

$$D(\boldsymbol{X},\boldsymbol{Y}) = \begin{cases} \dfrac{D_{LI}(\bar{\boldsymbol{X}},\bar{\boldsymbol{Y}})}{2m} & (m \geq n) \\ \dfrac{D_{LI}(\bar{\boldsymbol{X}},\bar{\boldsymbol{Y}})}{2n} & (m < n) \end{cases}$$

where $D_{LI}(\bar{\boldsymbol{X}},\bar{\boldsymbol{Y}})$ is the distance between $\bar{\boldsymbol{X}}$and $\bar{\boldsymbol{Y}}$.

In the multi-sample case, we used the median as the distance. Let $\boldsymbol{Y} = \{\boldsymbol{Y}_1,\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_k,\cdots,\boldsymbol{Y}_p\}$be multiple template signals. This distance was calculated as

$$D(\boldsymbol{X},\boldsymbol{Y}) = \underset{k=1,2,\cdots,p}{median}\left(D(\boldsymbol{X},\boldsymbol{Y}_k)\right)$$

where $D(\boldsymbol{X},\boldsymbol{Y}_k)$ is the normalized distance between an input signal and $k$-th template signals of multiple template signals.

### 3.3.1 Angular Velocity Distance

It is known that angular velocity does not depend on distance from the center of rotation. We calculate the absolute distance between the input signal and template signals. Even if signals of the same subject are selected, they do not correspond to the amplitude value from a difference in walking speed. To reduce differences between signals of the same subject, we normalized the signals by dividing the amplitude of each time by specific values. We adopted the method of normalization that divides amplitude of signal by the root mean square (RMS). The reason for using RMS for normalization is that it provided the best accuracy among some normalized methods in a preliminary experiment.

Let $\boldsymbol{g}_q = \left(g_q(1), g_q(2), \cdots, g_q(i), \cdots, g_q(m)\right)$ be the q-axis input angular velocity signal, and let $\boldsymbol{g}'_{qk} = \left(g'_{qk}(1), g'_{qk}(2), \cdots, g'_{qk}(j), \cdots g'_{qk}(n)\right)$ be the $k$-th template in q-axis angular velocity signal. We tested three distance calculation methods. They were Euclidean distance, Manhattan distance (absolute distance), and the distance based on correlation (denoted as Crr distance).

In the case of linear interpolation, Manhattan distance between a q-axis input angular velocity signal and a $k$-th template in q-axis angular velocity signal was calculated as

$$D_{LI}\left(\bar{\boldsymbol{g}}_q, \bar{\boldsymbol{g}}'_{qk}\right) = \sum_l \left|\bar{g}_q(l) - \bar{g}'_{qk}(l)\right|$$

Euclidean distance was calculated as

$$D_{LI}\left(\bar{\boldsymbol{g}}_q, \bar{\boldsymbol{g}}'_{qk}\right) = \sqrt{\sum_l \left(\bar{g}_q(l) - \bar{g}'_{qk}(l)\right)^2}$$

Crr distance was not divided by the total length of two signals, because signal length is not affected on the correlation value. Therefore, in the case of Crr distance calculation, $D\left(\boldsymbol{g}_q, \boldsymbol{g}'_{qk}\right)$ was calculated as

$$D\left(\boldsymbol{g}_q, \boldsymbol{g}'_{qk}\right) = 1 - NCC\left(\bar{\boldsymbol{g}}_q, \bar{\boldsymbol{g}}'_{qk}\right)$$

where $NCC(\,)$ is the normalized cross correlation between $\bar{\boldsymbol{g}}_q$ and $\bar{\boldsymbol{g}}'_{qk}$.

In the case of distance calculation using DTW, Euclidean distance calculation function $dist(g_q(i), g'_{qk}(j))$ was calculated as

$$dist(g_q(i), g'_{qk}(j)) = \left(g_q(i) - g'_{qk}(j)\right)^2$$

Finally, we calculated Euclidean distance with DTW as

$$D_{DTW}(\boldsymbol{g}_q, \boldsymbol{g}'_{qk}) = \sqrt{f(m,n)}$$

Manhattan distance calculation function was calculated as

$$dist(g_q(i), \acute{g}_{qk}(j),) = |g_q(i) - \acute{g}_{qk}(j)|$$

Crr distance cannot be calculated using DTW. Thus, in the case of DTW, we calculated only two types of distance.

### 3.3.2 Acceleration Distance

It is known that acceleration depends on the distance from the center of rotation in the circular motion. If different amplitude normalizations are applied to each axis acceleration, they are compressed at different ratios at the same time. As a result, when the normalized accelerations of the three axes at the same time were combined as a vector, the direction of the vector was changed before normalization. This problem was caused by comparing it with the values of acceleration. Hence, we compared it with the direction of three-axis acceleration between the input and the template acceleration signals [12]. Let $\boldsymbol{a}(i)$ be the $i^{\text{th}}$ input acceleration vector of an input signal, and let $\boldsymbol{a}'_k(j)$ be the $j^{\text{th}}$ template acceleration vector of a $k$ template signal.

$$\boldsymbol{a}(i) = \left(a_x(i), a_y(i), a_z(i)\right)$$
$$\boldsymbol{a}'_k(j) = \left(a'_{xk}(j), a'_{yk}(j), a'_{zk}(j)\right)$$

where $a_q(i)$ is the q-axis $i^{\text{th}}$ amplitude of input acceleration signal, and $a'_{qk}(j)$ is the q-axis $j^{\text{th}}$ amplitude of $k$-th template acceleration signal. The difference of direction between the $i^{\text{th}}$ input acceleration vector and $j^{\text{th}}$ $k$-th template acceleration vector was calculated as

$$dist(\boldsymbol{a}(i), \boldsymbol{a}'_k(j)) = \arccos \frac{\langle \boldsymbol{a}(i), \boldsymbol{a}'_k(j) \rangle}{\|\boldsymbol{a}(i)\| \|\boldsymbol{a}'_k(j)\|}$$

The distance based on difference of direction is calculated using DTW by substituting this function into $dist(\,)$.

To compare this three-axis composite method with others, authentication accuracy of each axis acceleration was calculated using the same normalization and distance calculation method for angular velocity.

## 3.4 Distance Fusion

In biometrics authentication, score fusions were attempted using various methods [1], [13].

To eliminate subject dependency, we subtracted the average distance from the distance before fusion. This average distance was calculated between a subject's template signal $\boldsymbol{Y}$ and the same subject's recorded data $\boldsymbol{\gamma}$ except his or her template signal $\boldsymbol{Y}$. This averaged distance was calculated with the all template signals selected in the evaluation. In the case of multi-sampling, the average distance of template signals $\overline{D(\boldsymbol{\gamma}, \boldsymbol{Y})}$ is an averaged distance of medians calculated among all combinations of template signals using the same calculation of the median between an input signal and multiple template signals. This averaged value was calculated as one value for a subject by using the subject's all selected template signals with replacing the template signals from subject's recorded signal. The normalized distance is calculated by subtracting the average distance from the distance calculated by DTW between an input signal and the template signals as

$$D_s(\boldsymbol{X}, \boldsymbol{Y}) = D(\boldsymbol{X}, \boldsymbol{Y}) - \overline{D(\boldsymbol{\gamma}, \boldsymbol{Y})}$$

Finally, two types of fused distance $D_f$ were calculated as

$$D_f = f\left(D_s(\boldsymbol{a}, \boldsymbol{a}'), D_s(\boldsymbol{g}_x, \acute{\boldsymbol{g}}_x), D_s(\boldsymbol{g}_y, \acute{\boldsymbol{g}}_y), D_s(\boldsymbol{g}_z, \acute{\boldsymbol{g}}_z)\right)$$

$$D_f = f(D_s(\boldsymbol{a}, \boldsymbol{a}'), D_s(\boldsymbol{a}_x, \boldsymbol{a}'_x), D_s(\boldsymbol{a}_y, \boldsymbol{a}'_y), D_s(\boldsymbol{a}_z, \boldsymbol{a}'_z), D_s(\boldsymbol{g}_x, \acute{\boldsymbol{g}}_x), D_s(\boldsymbol{g}_y, \acute{\boldsymbol{g}}_y), D_s(\boldsymbol{g}_z, \acute{\boldsymbol{g}}_z))$$

where $f(\,)$ is a function of fusion which combines the distances. In the our previous experiment, we found that the authentication from the four distances (difference of the direction between acceleration vectors and each axis angular velocity signal) outperformed the authentication from six distances (each axis acceleration signal and each axis angular velocity signal). Therefore, the accuracy from six distances was not verified in this study.

We consider four rules for fusing distances as below.

(1) Addition without weight coefficients (denoted as Sum)
(2) Support vector machine (SVM) with a linear kernel (denoted as Linear)
(3) SVM with a radial basis function kernel (denoted as RBF)
(4) Linear logistic regression (denoted as LLR)

In calculations using SVM, the classifier must learn based on training dataset. In this study, we obtained too many negative instances as compared with positive instances. Such datasets are said to be imbalanced. It is well known that SVM performs poorly in this case. Hence, we applied the synthetic minority over-sampling technique (SMOTE) [14] to adjust the number of these instances. This algorithm adds the instances among instances based on the $k$-nearest neighbor algorithm for the small instances.

# 4 EXPERIMENT

## 4.1 Index of Performance

We evaluated accuracy by equal error rate (EER). The EER is obtained from the intersection of the false acceptance rate (FAR) and the false rejection rate (FRR). An example of EER is shown in Fig. 12.

## 4.2 Dataset

Data was collected from 50 subjects, ranging in age from 18 to 21 years old. We instructed the subjects to walk at their normal walking speeds. When the measurement began, the subjects remained stationary for a few seconds. After that, they walked a specified distance once. The measurement course is a flat and straight indoor passageway. The subjects did not use a clock or metronome to measure their walking speed.

We set the sampling frequency of the sensor unit to 1,000 Hz. To equalize the performance of the smartphone's sensors, we changed the sampling frequency from 1,000 to 100 Hz by thinning out.

We obtained 30 signals of each axis acceleration and 30 signals of each axis angular velocity from every subject.

## 4.3 Experiment for Distance Calculation

Proposed authentication method is combined multiple processes. Verifying the all combinations of all methods of all processes needs too many times. Therefore, we applied step-by-step evaluation to find the appropriate method in each process.

To reveal the most appropriate distance calculation methods for each sensor, the authentication accuracy in uni-sensor uni-sample authentication was calculated by each distance calculation method. Table 2, and 3 show the EERs for uni-sensor uni-sample methods. The distance calculation combinations of DTW and Manhattan distance showed the best performance with each axis signal of two sensors. Therefore, this combination was employed as features of fusion functions.

## 4.4 Experiment for Verification of Fusion Effectiveness

To verify the effectiveness of each proposed method, we evaluate the four combinations. These are uni-sensor uni-sample combination, uni-sensor uni-sample combination, multi-sensor uni-sample combination, and multi-sample multi-sensor combination. Uni-sensor used one axis of a sensor. Multi-sensor calculated a fused score from six axes signals. Uni-sample used one template signal. Multi-sample used six template signals.

### 4.4.1 Experimental Setting

We divided the signals into five groups and performed five-fold cross-validation. To generate a fusion model, we used four groups as training data, and one group as test data. We calculated the distances between all of the training signals of all subjects. The distances between the same subjects are positive instances, and the distances between different

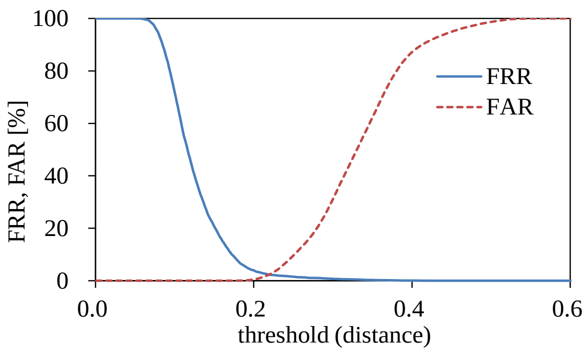



Figure 12: Determining EER.

subjects are treated as negative instances. The overall accuracies were calculated with common thresholds to each classifier in each fusion rule.

Template signals used for calculating distance include six signals, because the number of template signals is equal to the number of template signals of the previous study [7]. The manner of selecting templates from training data was to select six sequential signals from 24 signals. However, when some of the sequential six signals were selected as test data by cross-validation, we selected the signals in sequence from the nearest start time in the training data.

### 4.4.2 Experimental Result

For comparison purposes, we calculated EERs of uni-sensor uni-sample (one-axis and one template), uni-sensor multi-sample (one-axis and six templates), multi-sensor uni-sample (six-axis fusion and one template), multi-sensor multi-sample (six-axis fusion and six templates), and previous work [7]. The method used in this previous work created an authentication signal from six signals that were normalized for time length. They were calculated as the absolute distance between a template signal and an input signal.

We summarized the EERs in Tables 4, 5, and 6. The minimum EER (the best result) was 1.0%, which was achieved by the proposed multi-sensor multi-sample method from four distances with two SVMs. The best EER from the template generation and the distance calculation method of the previous work [6] to each axis signal for this dataset was 7.8%. Proposed method could outperform the previous study method.

Figures 13, 14, 15, and 16 show the receiver operating characteristics (ROC) curves for uni-sensor acceleration, uni-sensor angular velocity, multi-sensor fusion from four distances, and multi-sensor fusion from seven distances. These graphs show trade-off relations between the FAR and FRR. To compare the performance of each multi-sensor method, we plotted the ROC curves whose methods showed the best EER in the each combination in Fig. 17. From this ROC curve, we can observe that authentication method multi-sensor multi-sample with 4 distances shows the best performance.

Table 2: EERs [%] by each distance calculation method with DTW.

| | Manhattan distance | Euclidean distance |
|---|---|---|
| $\boldsymbol{a}_x$ | 8.8 | 10.4 |
| $\boldsymbol{a}_y$ | 5.3 | 6.0 |
| $\boldsymbol{a}$ | 4.6 | 6.2 |
| $\boldsymbol{g}_x$ | 6.6 | 7.0 |
| $\boldsymbol{g}_y$ | 8.2 | 9.7 |
| $\boldsymbol{g}$ | 7.4 | 8.9 |

Table 3: EERs [%] by each distance calculation method with linear interpolation.

| | Manhattan distance | Euclidean distance | Crr distance |
|---|---|---|---|
| $\boldsymbol{a}_x$ | 14.0 | 20.5 | 19.2 |
| $\boldsymbol{a}_y$ | 11.2 | 14.1 | 7.9 |
| $\boldsymbol{a}$ | 13.4 | 18.3 | 11.6 |
| $\boldsymbol{g}_x$ | 14.1 | 15.9 | 11.2 |
| $\boldsymbol{g}_y$ | 15.0 | 18.9 | 16.3 |
| $\boldsymbol{g}$ | 15.2 | 19.1 | 9.5 |

Table 4: Uni-sensor EERs [%].

| | Uni-sensor uni-sample authentication | Uni-sensor multi-sample authentication |
|---|---|---|
| $\boldsymbol{a}_x$ | 8.8 | 4.5 |
| $\boldsymbol{a}_y$ | 5.3 | 2.2 |
| $\boldsymbol{a}$ | 4.6 | 2.2 |
| $\boldsymbol{g}_x$ | 6.6 | 2.4 |
| $\boldsymbol{g}_y$ | 8.2 | 3.1 |
| $\boldsymbol{g}$ | 7.4 | 3.0 |

Table 5: Multi-sensor EERs [%] fused four distances.

| | Multi-sensor uni-sample authentication | Multi-sensor multi-sample authentication |
|---|---|---|
| Sum | 1.7 | 1.2 |
| Linear | 1.5 | 1.0 |
| RBF | 1.4 | 1.0 |
| LLR | 1.5 | 1.1 |

Table 6: Multi-sensor EERs [%] fused seven distances.

| | Multi-sensor uni-sample authentication | Multi-sensor multi-sample authentication |
|---|---|---|
| Sum | 1.5 | 1.1 |
| Linear | 1.4 | 1.2 |
| RBF | 1.2 | 1.1 |
| LLR | 1.5 | 1.3 |

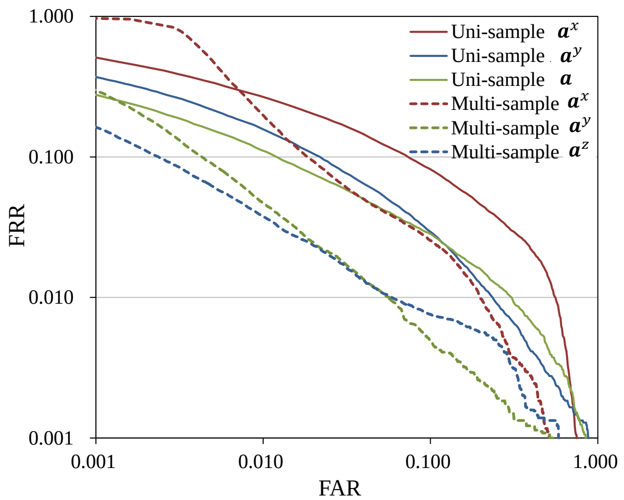



Figure 13: ROC curves of uni-sensor method for three-axis acceleration.

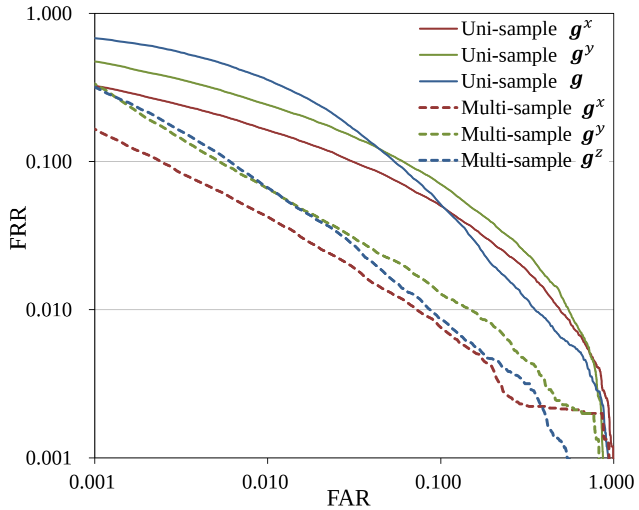



Figure 14: ROC curves of uni-sensor methods for three-axis angular velocity.

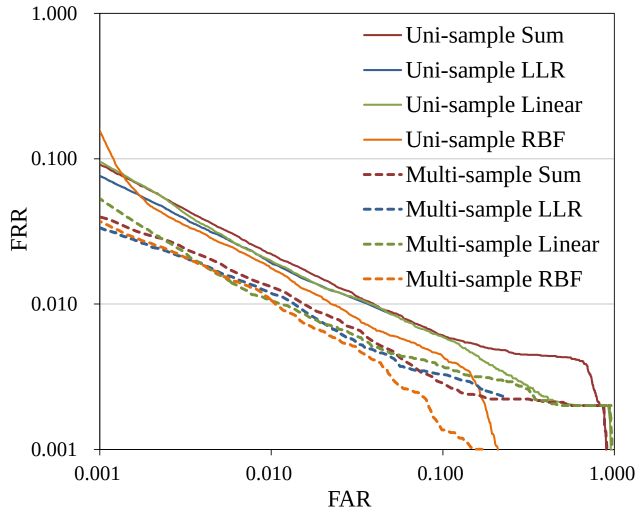



Figure 15: ROC curves of multi-sensor methods from four distances.

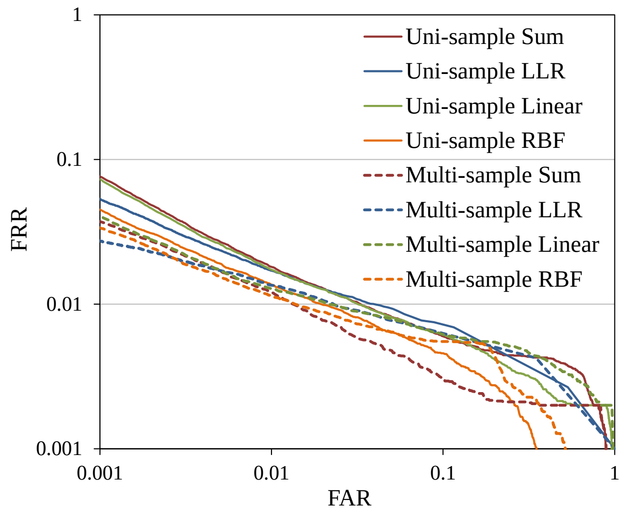



Figure 16: ROC curves of multi-sensor methods from seven distances.

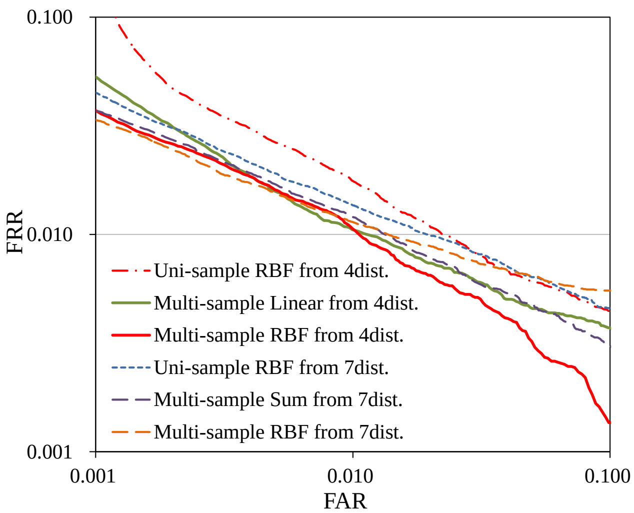



Figure 17: ROC curves of the best EER methods in each multi-sensor combination.

## 5 DISCUSSION

This research described here is an effort to improve the accuracy of gait-based authentication. We applied multi-sensor and multi-sample fusion to improve accuracy as compared with the uni-sensor and uni-sample method. The results of the tests are summarized in Tables 4, 5, and 6.

Comparing the effect of uni-sample and multi-sample, uni-sensor multi-sample (the best EER = 2.2%) is smaller than uni-sensor uni-sample (the best EER = 4.6%), giving a relative reduction in EER of 52.2%. Similarly, the multi-sensor multi-sample method from four distances (the best EER = 1.0%) is smaller than multi-sensor uni-sample method (the best EER = 1.4%), giving a relative reduction in EER of 28.6%. The multi-sensor multi-sample method from the seven distances (the best EER = 1.1%) is smaller than multi-sensor uni-sample method (the best EER = 1.2%), giving a relative reduction in EER of 8.3%. The results show that the multi-sample method is effective in gait-based authentication. However, in the case of multi-sensor authentication from the seven distances, the effect of multi-sample is smaller than the case of the other authentications.

Comparing the effect of multi-sensor, multi-sensor uni-sample from four distances is smaller than uni-sensor uni-sample, with a relative reduction of 69.9% in EER. EER of multi-sensor multi-sample from the four distances is smaller than EER of uni-sensor multi-sample, with a relative reduction in EER of 54.5%. Similarly, EER of multi-sensor uni-sample from seven distances is smaller than EER of uni-sensor uni-sample, with a relative reduction in EER of 73.9%. EER of multi-sensor multi-sample from seven distances is smaller than EER of uni-sensor multi-sample, with a relative reduction in EER of 45.5%. The results show that the multi-sensor method is effective in gait-based authentication.

Comparing the effect of fusion algorithms, SVM with RBF indicated the best performance in the four types of multi-sensor authentication. Hence, SVM with RBF is effective in gait-based authentication.

Comparing the effect of the multiple distances, in the uni-sample authentication, the best EER from seven distances show the smaller value than the best EER from four distances. However, the best EER of four distances show the smaller value than the best EER from seven distances in the multi-sample authentication. We think that Manhattan distances from three-axis acceleration signals ineffective for accuracy improvement in comparison with difference of direction of three-axis acceleration vectors. This is for the following reasons. EER from four distances shows almost same value of EER from seven distances. SVM with RBF and multi-sample from seven distances become significantly lower performance than SVM with RBF and multi-sample from four distances in the low FRR area.

Table 7 summarizes the best EERs of previous work with a uni-sensor. The conditions are different from those of our study, making simple comparison of the results of all the work difficult. However, we outperformed all of the previous methods. The work in [6] authenticated 50 subjects whose acceleration sensor was in a front trouser pocket. The condition of sensor location was similar to that of our experiment. The best result (EER = 7.3%) was obtained by calculating the Manhattan distance. The proposed method outperforms this previous work under the same sensor condition. Furthermore, our proposed method also outperforms the distance calculation and template generation methods of previous work on the same dataset.

In this experiment, to obtain the subjects' signals without effect of pocket form, the sensor unit was attached to their thigh. From the result of this experiment, we evaluated the accuracy based on the pure signals without noise caused by the pocket. However, when considering the case of actual use, we need to evaluate the accuracy of the proposed method under the condition of putting the sensors into subjects' pockets.

This study was researched based on the premise that walking patterns of all people were the feature which could identify only one person. This premise is not clear. However, the researches of gait recognition based on dynamic image achieved high accuracy with large number of the subjects. According to these researches, there is possibility that walking pattern is unique feature. Furthermore, it is not known whether the distribution of features from walking signals is uniform. It is important to clarify these issues for updating this research. To verify these things, we will need to collect the large scale dataset.

## 6 CONCLUSION

This paper proposed multi-sample and multi-sensor method for accuracy improvement of gait-based authentication and verified the effectiveness of the proposed methods.

First, we observed the relation among the steps and six-axis signals in order to extract the quasi-periodic signals generated by walking motion of the same gait phase order in all subjects by using two force sensors and a six-axis sensor. These findings show that it is possible to divide into quasi-periodic signals by extracting x-axis acceleration from local maxima to next local maxima.

Table 7: EERs of uni-sensor-based authentication work.

| Reference | Number of subjects | Best EER[%] |
|---|---|---|
| Jani et al. [4] | 36 | 7 |
| Gafurov et al. [5] | 22 | 16 |
| Gafurov et al. [6] | 21 | 5 |
| Gafurov et al. [7] | 50 | 7.3 |
| Gracian et al. [8] | 11 | 3 |
| Derawi et al. [9] | 60 | 5.7 |
| This paper | 50 | 1.0 |

Next, to improve the authentication accuracy, we proposed multi-sensor and multi-sample authentication methods. To find the appropriate methods for each process, we verified the authentication accuracy calculated by each method, distance calculation methods, fusion algorithms, multi-sensor, and multi-sample. From these experiments, we could find the appropriate methods to obtain the better authentication accuracy in the gait-based authentication.

We evaluated the proposed method with 50 subjects. The best EER performance was 1.0%, which was achieved by the combination of multi-sensor multi-sample using SVM with RBF from four distances. These results show that multi-sensor multi-sample authentication is useful for gait-based authentication. We confirmed that the proposed method using appropriate methods which were obtained from this study leads to better performance than the conventional methods.

In the future, we need to collect the gait data in an experimental condition that is similar to an actual use environment (e.g., a corner, a slope, with different type of pockets, and with different type of shoes) for the feasibility experiments. Furthermore, this experiment could indicate the combination effect of acceleration-based authentication to angular velocity-based authentication. However, it was not clear whether the combination of angular velocity-based authentication to acceleration-based authentication is effective. We think we will try to evaluate this question under the actual use environment, after collecting the large dataset.

## REFERENCES

[1] A.K. Jain, A. Ross, and S. Prabhakar, “An Introduction to Biometric Recognition,” IEEE Trans. Circuits and Systems for Video, Vol.14, No.1, pp.4-20 (2004).

[2] F. Alonso-Fernandez, J. Fierrez, D. Ramos, and J. Ortega-Garcia, “Dealing with sensor interoperability in multi-biometrics: The UPM experience at the Biosecure multimodal Evaluation 2007,” Proc. of SPIE Defense and Security Symposium, Workshop on Biometric Technology for Human Identification, pp.69440J-69440J (2008).

[3] X. Zhou and B. Bhanu, “Feature Fusion of Side Face and Gait for Video-based Human Identification,” Pattern Recognition, Vol.41, pp.778-795 (2008).

[4] J. Mäntyjärvi, M. Lindholm, E. Vildjiounaite, S. Mäkelä, and H.A. Ailisto, “Identifying Users of Portable Devices from Gait Pattern with Accelerometers,” Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing, pp.973-976 (2005).

[5] D. Gafurov, E. Snekkenes, and T.E. Buvarp, “Robustness of Biometric Gait Authentication Against Impersonation Attack,” Proc. of the 1st International Workshop on Information Security (IS'06), OnTheMove Federated Conferences (OTM'06), Vol.4277, pp.479-488 (2006).

[6] D. Gafurov, K. Helkala, and T. Sondrol, “Biometric Gait Authentication Using Accelerometer Sensor,” Journal of Computers, Vol.1, No.7, pp.51-59 (2006).

[7] D. Gafurov, E. Snekkenes, and P. Bours, “Gait Authentication and Identification using Wearable Accelerometer sensor,” Proc. of IEEE Workshop on Automatic Identification Advanced Technologies, pp.220-225 (2007).

[8] G. Trivino, A. Alvarez-Alvarez, and G. Bailador, “Application of the Computational Theory of Perceptions to Human Gait Pattern Recognition,” Pattern Recognition, Vol.43, pp.2572-2581 (2010).

[9] M.O. Derawi, P. Bours, and K. Holien, “Improved Cycle Detection for Accelerometer Based Gait Authentication,” Proc. of the 6th International Conference on Intelligent Information Hiding and Multimedia Signal Processing(IIH-MSP '10), pp.312-317 (2010).

[10] S. Mondal, A. Nandy, P. Chakraborty, and G.C. Nandi, “Gait Based Personal Identification System Using Rotation Sensor,” Journal of Emerging Trends in Computing and Information Sciences, Vol.3, No.3, pp.395-402 (2012).

[11] A. Savitzky and M.J.E. Golay, “Smoothing and Differentiation of Data by Simplified Least-squares Procedures,” Anal. Chem., Vol.36, No.8, pp.1627-1639 (1964).

[12] F. Okumura, A. Kubota, Y. Hatori, K. Matsuo, M. Hashimoto, and A. Koike, “A Study on Biometric Authentication based on Arm Sweep Action with Acceleration Sensor,” Proc. of International Symposium on Intelligent Signal Processing and Communications, 2006(ISPACS '06), pp.219-222 (2006).

[13] S. BenYacoub, “Multi-Modal Data Fusion for Person Authentication using SVM,” Proc. 2nd Int. Conf. Audio-Video Based Biometric Person Authentication, pp.25-30 (1999).

[14] N.V. Chawla, K.W. Bowyer, L.O. Hall, and W.P. Kegelmeyer, “SMOTE:Synthetic Minority Over-sampling Technique, Journal of Artificial Intelligence Research,” Vol.16, pp.321-357 (2002).

(Received October, 11, 2015)

**Shinsuke Konno** received B.E. and M.S. from Hokkaido University in 2002 and 2004, respectively. He is currently a research associate at the National Institute of Technology, Hakodate Collage. His current research interests include mobile sensing and biometrics. He is a member of IEICE, and IPSJ.

**Yoshitaka Nakamura** received B.E., M.S., and Ph.D. degrees from Osaka University in 2002, 2004 and 2007, respectively. He is currently a research associate at the School of Systems Information Science, Future University Hakodate. He is a member of IEEE, IEICE, and IPSJ.

**Yoh Shiraishi** received doctor's degree from Keio University in 2004. He is currently an associate professor at the Department of Media Architecture, School of Systems Information Science, Future University Hakodate, Japan. His research interests include database, mobile sensing and ubiquitous computing. He is a member of IPSJ, IEICE, GISA and ACM.

**Osamu Takahashi** received master's degree from Hokkaido University in 1975. He is currently a professor at the Department of System Information Science at Future University Hakodate, Japan. His research interest includes ad-hoc network, network security, and mobile computing. He is a member of IEEE, IEICE.

# PBIL-RS: Effective Repeated Search in PBIL-based Bayesian Network Learning

Yuma Yamanaka†*, Takatoshi Fujiki†, Sho Fukuda†, and Takuya Yoshihiro‡**

†Graduate School of Systems Engineering, Wakayama University, Japan
‡Faculty of Systems Engineering, Wakayama University, Japan
* s161065@sys.wakayama-u.ac.jp
** tac@sys.wakayama-u.ac.jp

***Abstract*** - Bayesian Network, which is a model of probabilistic causal relationship, is an practically important graphical model learned from observation data. To learn a near-optimal Bayesian network model from a set of observation data, efficient optimization algorithm is required to search an exponentially large solution space, as this problem was proved to be NP-hard. To find better Bayesian network models in limited time, several efficient approximated search algorithms have been proposed such as genetic algorithms. Among them, algorithms based on probability vectors such as PBIL (Population-Based Incremental Learning) are regarded as a better sort of algorithms to learn superior Bayesian networks in a practical computation time. However, PBIL has a problem that it finishes executing when it converges. Namely, after convergence, it cannot find any better solutions. To solve the problem, in this paper, we propose PBIL-RS (PBIL-Repeated Search), which is an improvement of PBIL. In PBIL-RS, if the search area becomes sufficiently small in the process of converging the probability vector, we in turn spread the search area and again begin the converging process, repeatedly. We performed an evaluation of PBIL-RS and showed that it outperforms the existing algorithms in the PBIL-family. We further explored the behavior of PBIL-RS and found several key behaviors that lead the characteristics to find better solutions.

***Keywords***: Bayesian Networks, PBIL, Evolutionary Algorithm, EDA, Information Criterion

## 1 INTRODUCTION

Bayesian Network is used as a probabilistic model to analyze causal relationship between events from data. Recently, rapid growth of the Internet and processing speed of computers have made us possible to analyze the causal relationship from a large amount of data, and Bayesian Network is one of the important data analysis methods that are useful in various research fields with large data such as bioinformatics, medical analyses, document classifications, information searches, decision support, etc.

However, there is one difficulty that learning Bayesian Network models is proved to be NP-hard [1]. In other words, solution space exponentially increases as the number of events in the Bayesian Network increases. Therefore, several near-optimal algorithms to find better Bayesian Network models within a limited time have been proposed so far. Cooper et al. proposed an algorithm to learn Bayesian Networks called K2 that reduced execution time by limiting the search space [2]. To limit the search space, K2 applies a constraint in the order of events. The order constraint, for example, means that future events cannot be caused of events in the past. However, in many practical cases, we cannot assume such an order constraint. Therefore, to learn Bayesian Networks in general cases, several approaches have been proposed. Many of them use genetic algorithms (GAs), which find better Bayesian Network models when we take more time for computation [3]-[5]. Meanwhile, recently, requirements for large-data analyses arise due to the growth of the Internet. To meet these requirements, more efficient algorithms to find better Bayesian Network models within smaller time are strongly expected.

On the background above, a number of authors have proposed a new category of algorithms. Those algorithms called EDA (Estimation of Distribution Algorithm) have been reported to find better Bayesian Network models [6]-[8]. EDA is a kind of genetic algorithms that evolves statistic distributions from which we produce individuals over generations. Namely, EDA is a stochastic optimization algorithm. From the result of Kim et al., PBIL-based algorithm performed the best among several EDA-based algorithms [7].

Blanco et al. presented the first PBIL-based algorithm for Bayesian Networks [9]. They showed that their PBIL-based algorithm outperforms the traditional K2 algorithms. However, his algorithm has a drawback that his algorithm easily falls into local minimum solutions because it does not include any mutation operation to avoid converging into local minimum solutions. To overcome this drawback, several mutation operations were proposed for PBIL-based algorithms to learn Bayesian Networks. Handa et al. introduced bitwise mutation (BM), which apply mutation operations in which each edge is added or deleted with a constant probability [6]. Kim et al. proposed transpose mutation (TM) that is designed specific to Bayesian Networks [7]. This operation changes the direction of edges in the individuals produced in each generation. Fukuda et al. proposed a mutation operator called probability mutation (PM) for PBIL-based algorithms to learn Bayesian Networks [8]. Probability mutation manipulates the probability vector to avoid converging at local minimum solutions. These mutation operators improved the performance of PBIL-based algorithms to learn Bayesian Networks by avoiding local minimum solutions. However, mutation operators also have a drawback that the searching area jumps to other areas due to mutations before searching the local areas deep enough to explore good solutions.

In this paper, we propose a new PBIL-based algorithm called PBIL-RS (PBIL-Repeated Search), which is an improvement of PBIL. In PBIL-RS, if the search area becomes sufficiently small in the process of converging the probability vector, we in turn spread the search area and again begin the converging process, repeatedly. By searching local areas deeply until the probability vector converges into a sufficiently small area, PBIL-RS improves the performance of PBIL-based algorithms to learn Bayesian Networks. We performed an evaluation of PBIL-RS, clarified its characteristics, and showed the superiority in its performance.

The rest of this paper is organized as follows: In Section 2, we give the basic definitions on Bayesian Networks and also describe related work in this area of study. In Section 3, we propose a new efficient search algorithm called PBIL-RS to achieve better learning performance of Bayesian Networks. In Section 4, we describe the evaluation of PBIL-RS, and finally we conclude this paper in Section 5.

## 2 PRELIMINARY DEFINITIONS

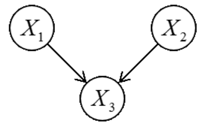
### 2.1 Bayesian Network

A Bayesian Network model visualizes the causal relationship among events through graph representation. In a Bayesian Network model, events are represented by nodes while causal relationships are represented by edges. See Fig. 1 for a concise example. Nodes $X_1, X_2$, and $X_3$ represent distinct events, where they take 1 if the corresponding events occur, and take 0 if the events do not occur. Edges $X_1 \to X_3$ and $X_2 \to X_3$ represent causal relationships, which mean that the probability of $X_3 = 1$ depends on events $X_1$ and $X_2$ . If edge $X_1 \to X_3$ exists, we call that $X_1$ is a parent of $X_3$ and $X_3$ is a child of $X_1$. Because nodes $X_1$ and $X_2$ do not have their parents, they have own prior probabilities $P(X_1)$ and $P(X_2)$. On the other hand, because node $X_3$ has two parents $X_1$ and $X_2$, it has a conditional probability $P(X_3|X_1, X_2)$. In this example, the probability that $X_3$ occurs is 0.890 under the assumption that both $X_1$ and $X_2$ occur. Note that, from this model, Bayesian inference is possible: if $X_3$ is known, then the posterior probability of $X_1$ and $X_2$ can be determined, which enables us to infer more accurately the occurrence of events.

The Bayesian Networks model can be learned from the data obtained through the observation of events. Let $O = \{o_j\}$, $(1 \le j \le S)$ be a set of observations, where $S$ is the number of observations. Let $o_j = (x_{j1}, x_{j2}, \ldots, x_{jN})$ be $j$-th observation, which is a set of observed values $x_{ji}$ on event $X_i$ for all $i(1 \le i \le N)$, where $N$ is the number of events. We try to learn a good Bayesian Network model $\theta$ from the given set of observations. Note that, good Bayesian Network model $\theta$ is the one that creates data sets similar to the original observation $O$. As an evaluation criterion to measure the level of fitting between $\theta$ and $O$, we use AIC (Akaike's Information Criterion) [10], which is one of the best known criterion used in Bayesian Networks. Formally, the problem of learning Bayesian Networks that we consider in this paper is defined as follows:

**Problem 1:** From the given set of observations $O$, compute a Bayesian Network model $\theta$ that has the lowest AIC criterion value.

| $X_1$ | $P(X_1)$ |
|---|---|
| 0 | 0.79 |
| 1 | 0.21 |




| $X_2$ | $P(X_2)$ |
|---|---|
| 0 | 0.88 |
| 1 | 0.12 |

| $X_1$ | $X_2$ | $P(X_3 \mid X_1, X_2)$ | |
|---|---|---|---|
| | | 0 | 1 |
| 0 | 0 | 0.999 | 0.001 |
| 0 | 1 | 0.710 | 0.290 |
| 1 | 0 | 0.360 | 0.640 |
| 1 | 1 | 0.110 | 0.890 |

Figure 1: A Bayesian Network Model

### 2.2 PBIL

Recently, a category of the evolutionary algorithms called EDA (Estimation Distribution Algorithm) appears and reported to be efficient to learn Bayesian Network models. As one of EDAs, PBIL was proposed by Baluja et al. in 1994, which is based on genetic algorithm designed to evolve a probability vector [13]. Later, Blanco et al. applied PBIL to the Bayesian Network learning, and showed that PBIL efficiently works in this problem [9]. In PBIL, an individual creature s is defined as a vector $s = \{v_1, v_2, \ldots, v_L\}$, where $v_i(1 \le i \le L)$ is the $i$-th element that takes a value 0 or 1, and $L$ is the number of elements that consist of an individual. Let $P = \{p_1, p_2, \ldots, p_L\}$ be a probability vector where $p_i(1 \le i \le L)$ represents the probability to be $v_i = 1$. The algorithm of PBIL is described as follows:

(1) As initialization, we let $p_i = 0.5$ for all $i = 1, 2, \ldots, L$.

(2) Generate a set $S$ that consists of $C$ individuals according to probability vector $P$, i.e., element $v_i$ of each individual is determined by the corresponding probability $p_i$.

(3) Compute the evaluation score for each individual $s \in S$ (In this paper we use AIC as the evaluation score).

(4) Select a set of individuals $S'$ whose members have evaluation scores within top $C'$ in $S$, and update the probability vector according to $S'$ . Specifically, the formula applied to every $p_i$ to update the probability vector is shown as follows.

$$p_i^{new} = ratio(i) \times \alpha + p_i \times (1 - \alpha), \tag{i}$$

where $p_i^{new}$ is the updated value of the new probability vector ($p_i$ is soon replaced with $p_i^{new}$), $ratio(i)$ is the function that represents the ratio of individuals in $S'$ that include edge $i$ (i.e., $v_i = 1$), and $\alpha$ is the parameter called learning ratio.

(5) Repeat steps (2)-(4) until $P$ converges.

By merging top-$C'$ individuals, PBIL evolves the probability vector such that the good individuals are more likely to be generated. Different from other genetic algorithms, PBIL does not include “crossover” between individuals. Instead, it evolves the probability vector as a “parent” of the generated individuals.

## 2.3 PBIL-based Bayesian Network Learning

In this section, we describes a PBIL-based algorithm that learns Bayesian Network models. Because our problem (i.e. Problem 1) to learn Bayesian Network models is a little different from the general description of PBIL shown in the previous section, a little adjustment is required. In our problem, individual creatures correspond to each Bayesian Network model. Namely, with the number of events $N$, an individual model is represented as $s = \{v_{11}, v_{12}, \ldots, v_{1N}, v_{21}, v_{22}, \ldots, v_{N1}, v_{N2}, \ldots, v_{NN}\}$ where $v_{ij}$ corresponds to the edge from an event $X_i$ to $X_j$, i.e., if $v_{ij} = 1$, the edge from $X_i$ to $X_j$ exists in $s$, and if $v_{ij} = 0$ it does not exist. Similarly, we have the probability vector $P$ to generate individual models as $P = \{p_{11}, p_{12}, \ldots, p_{1N}, p_{21}, p_{22}, \ldots, p_{N1}, p_{N2}, \ldots, p_{NN}\}$ where $p_{ij}$ is the probability that the edge from $X_i$ to $X_j$ exists. A probability vector can be regarded as a table as illustrated in Fig. 2. Note that, because Bayesian Networks do not allow self-edges, $p_{ij}$ is always 0 if $i = j$. The process of the proposed algorithm is basically obtained from the steps of PBIL, as described in the following.

(1) Initialize the probability vector $P$ as $p_{ij} = 0$ if $i = j$, and $p_{ij} = 0.5$ otherwise, for each $i, j (1 \leq i, j \leq N)$.

(2) Generate $S$ as a set of $C$ individual models according to $P$. (This step (2) is illustrated in Fig. 3)

(3) Compute the evaluation scores for all individual models $s \in S$.

(4) Select a set of individuals $S'$ whose members have top-$C'$ evaluation values in $S$, and update the probability vector according to the formula (i). (These steps (3) and (4) are illustrated in Fig. 4.)

(5) Repeat steps (2)-(4) until $P$ converges.

Same as PBIL, the proposed algorithm evolves the probability vector so that we can generate better individual models. However, there is a point specific to Bayesian Networks, that is, a Bayesian Network model is not allowed to have cycles in it. To consider this point in our algorithm, step 2 is detailed as follows:

(2a) Consider every pair of events $(i, j)$ where $1 \leq i, j \leq N$ and $i \neq j$, create a random order of them.

(2b) For each pair $(i, j)$ in the order created in step (2a), determine the value $v_{ij}$ according to $P$; every time $v_{ij}$ is determined, if $v_{ij}$ is determined as 1, we check whether this edge from $X_i$ to $X_j$ creates a cycle with all the edges determined to exist so far. If it creates a cycle, let $v_{ij}$ be 0.

(2c) Repeat steps (2a) and (2b) until all the pairs in the order are processed.

| **P** | | Parent Node | | | | | |
|---|---|---|---|---|---|---|---|
| | | $X_1$ | $X_2$ | $\ldots$ | $X_i$ | $\ldots$ | $X_N$ |
| Child Node | $X_1$ | 0.0 | 0.5 | $\ldots$ | $p_{i1}$ | $\ldots$ | 0.5 |
| | $X_2$ | 0.5 | 0.0 | $\ldots$ | $p_{i2}$ | $\ldots$ | 0.5 |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ | $\ldots$ | $\vdots$ |
| | $X_j$ | $p_{1j}$ | $p_{2j}$ | $\ldots$ | $p_{ij}$ | $\ldots$ | $p_{Nj}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| | $X_N$ | 0.5 | 0.5 | $\ldots$ | $p_{iN}$ | $\ldots$ | 0.0 |

Figure 2: A Probability Vector

These steps enable us to learn good Bayesian Network models within the framework of PBIL. Note that the algorithm introduced in this section does not include mutation operators. Therefore, naturally, it easily converges to a local minimum solution. To avoid converging to the local minimum solution and to improve the performance of the algorithm, several mutation operations have been proposed. A mutation operator called bitwise mutation (BM) was introduced by Handa [6]. BM applies mutations to each edge in each individual with a certain mutation probability. Kim et al. proposed a mutation operator called transpose mutation (TM), which is specifically designed for Bayesian Networks [7]. TM changes the direction of edges in the individuals produced in each generation. Fukuda et al. proposed a mutation operator called probability mutation (PM) for PBIL-based Bayesian Network learning [8]. PM manipulates the probability vector to avoid converging at local minimum solutions. These mutations avoid converging at local minimum solutions, and it improves the efficiency to learn Bayesian Networks with PBIL-based algorithms.

# 3 PROPOSED ALGORITHM: PBIL-RS

We propose PBIL-RS (PBIL- Repeated Search), which is an algorithm to learn Bayesian Networks based on PBIL. To search for good Bayesian Networks efficiently, we introduce a new technique instead of mutation operators. Because mutation operators work with a certain mutation probability, they tend to change the search space before we deeply search the current search area to explore good solutions. As a result, efficiency of the algorithm decreases by skipping the search areas where many superior solutions are likely to be buried. In contrast, in PBIL-RS, we transit the search space only after we search the current search area deeply, i.e., only after PBIL-RS judged that the search space gets converged. With this technique, we can search deeply the specific space in which superior solutions would exist while avoiding local minimums.

Figure 5 shows the outline of PBIL-RS. In general, in the search space of Bayesian Network models, there are many local minimum points. Because models with similar structures tend to have similar evaluation scores, superior solutions would likely be collected at several local areas in the search

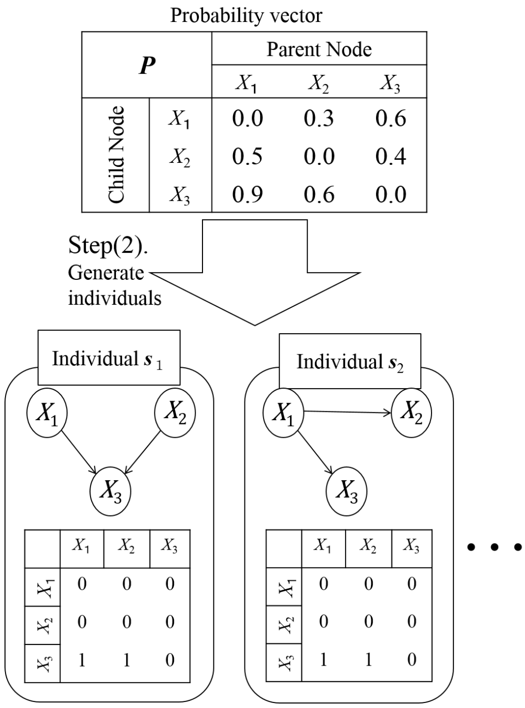



Figure 3: Generating Individuals from Probability Vector

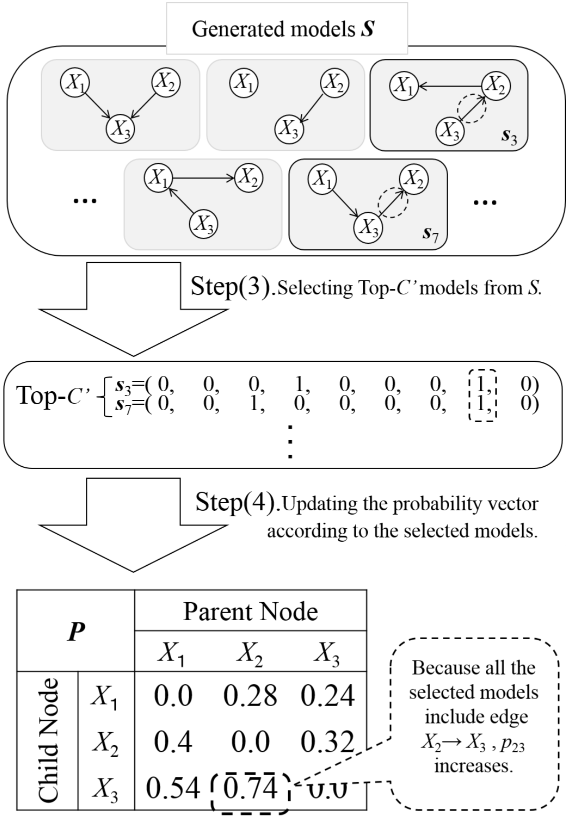



Figure 4: Step(3)(4): Updating Probability Vector

space. Our algorithm PBIL-RS explores these areas with the following steps: (1) Initially, PBIL-RS sets the search space as the whole solution space. (2) As the algorithm proceeds and the generation grows, the search space usually gets smaller by focusing on an area in which superior solutions would be likely to exist. When the search space converges to a sufficiently small area, and PBIL-RS judges that the current area is sufficiently searched out, and (3) PBIL-RS in turn spreads the search space to explore different local minimum areas. Here, if the size of the spread search space is not sufficiently large, it may again fall into the same local minimum area. In order to avoid this, PBIL-RS spreads it to be larger search spaces step-by-step. Specifically, the size of the spread search space is firstly small to search near local minimum areas, and if we cannot find superior solutions in the next convergence, we then try to spread to larger search spaces to reach more distant search areas.

PBIL-RS controls the search space with probability vector $P$. Each element $p_{i,j}$ of vector $P$ represents the probability to have the corresponding edge $(i, j)$ in the generated Bayesian Network models. Thus, if each element $p_{i,j}$ approaches to 0 or 1, then naturally we have a probabilistic bias in the structure of the generated Bayesian Network models: The closer to 0.5 each element of probability vector $P$ is, the larger the variation of generated models, and the closer to 0 or 1 each element is, the smaller the variation is. Namely, the probability vector $P$ controls the variation and the bias of the generated structures of Bayesian Network models. Based on this, for probability vector $P$, we define *convergence level* $S$ as follows:

$$S = \frac{\Sigma_{i,j(i \neq j)}\{0.5 - |P_{ij} - 0.5|\}}{N(N-1)}. \qquad \text{(ii)}$$

*Convergence level* $S$ takes the average of the difference between 0 (or 1) and each element of probability vector $P$. Namely, the less this value is, the smaller search space is. In PBIL-RS, generally *convergence level* $S$ gets smaller as generation proceeds. Thus, PBIL-RS spreads the search space when the search space shrinks to be sufficiently small. To judge that the search space is sufficiently small, we introduce the number of search limitation $k$. Specifically, when *convergence level* $S$ does not update the smallest value in the past $k$ generations, i.e., the *convergence level* $S$ in the $k$-th last generation takes the smallest value in the past $k$ generations, PBIL-RS judges that the search space is sufficiently small and has converged.

When PBIL-RS detects the search space convergence, it in turn spreads the search space. We define $H$ as the level to spread the search space. PBIL-RS modifies the probability vector $P$ to increase the *convergence level* $S$ to $H$. Specifically, we choose an element of $P$ randomly, and reset it as $P_{ij} = 0.5$. This operation repeats until $S \leq H$ holds.

In addition, as mentioned previously, we change the value $H$ dynamically to spread the search space and so avoid converging to the same local minimum areas repeatedly. More specifically, (a) we firstly initialize $H$ with the initial value $H_{min}$, (b) secondly every time the search space is converged we increase $H$ by a constant value *spread width* $H_{inc}$, and (c) lastly when we find the solution that has the best score so far, we again initialize $H$ with the initial value. This operation enables PBIL-RS to leave a local area quickly when good solutions would hardly be found, and guide to the bigger search

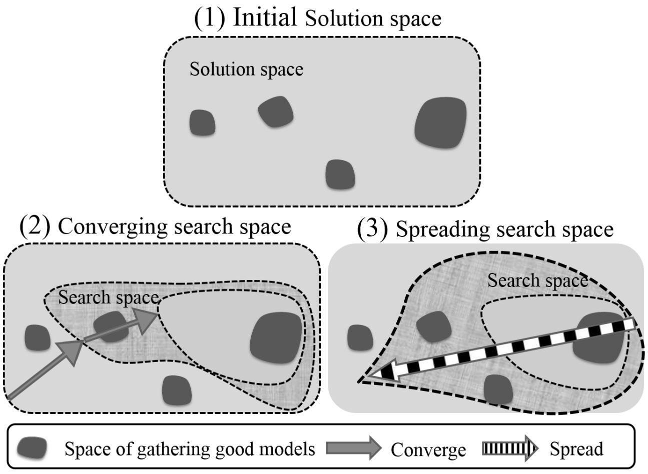



Figure 5: PBIL-RS Method

space.

The formal description of PBIL-RS is as follows. Processes (i)-(iv) are inserted into the steps (4) and (5) described in subsection 2.3.

(i) If the Bayesian Network model that has the best score so far is found, $H$ is initialized by the initial value $H_{min}$.

(ii) Choose an element $p_{ij}$ in $P$ randomly, and reset it as $p_{ij} = 0.5$.

(iii) If $S < H$, then return to step (ii).

(iv) Augment $H$ by *spread width* $H_{inc}$.

# 4 EVALUATION

## 4.1 Purpose of Evaluations

We evaluate PBIL-RS to measure the performance and also to explore the key behavior of PBIL-RS to find better solutions. In our evaluation, we first clarify that PBIL-RS performs well even if the number of events vary, and next we investigate the effect of several parameter variables on the performance. Clarifying the good parameter values would be helpful for us in determining the parameter values in the practical scenes. Through these evaluation results, we would find that the behavior of PBIL-RS is favorable to find better Bayesian Network models efficiently.

## 4.2 Performance Comparison with Existing Methods

We designed our evaluation procedure as follows: We select Bayesian Network models used in our evaluation. In this paper, we use two well-known Bayesian Network models called Alarm Network [11] and Pathfinder [12], where Alarm Network represents the causal relation among events to monitor patients in intensive care units, and Pathfinder represents that related to the diagnosis of lymph node diseases. Note that Alarm Network includes 37 nodes and Pathfinder does 135 nodes. We generate an observation data set from each of the two models. In a Bayesian Network model, each node has a set of conditional probability so that we can obtain a set of values corresponding to all nodes according to the probability. Figure 6 shows an example of the data set generated from the example of Bayesian Networks shown in Fig 1, where $j$-th row represents an example of $j$-th observation set $o_j$ generated according to the conditional probabilities of the model. We generated a data set that consists of 1,000 observations from each of two models. We use AIC criterion as the evaluation score, which is one of the representative criterion to measure the distance between the input data set and a Bayesian Network model.

Data set

| Observation No. | $X_1$ | $X_2$ | $X_3$ |
|---|---|---|---|
| 1 | 0 | 1 | 0 |
| 2 | 1 | 0 | 1 |
| 3 | 1 | 1 | 1 |
| ⋮ | ⋮ | ⋮ | ⋮ |

Repeatedly determine values of all nodes to generate a data set

Figure 6: Generating an Observation Data Set

Table 1: Parameters of PBIL-RS

| **Parameters** | **Values** |
|---|---|
| # of observations | 1000 |
| Individuals in a generation ($C$) | 1000 |
| # of selected individuals ($C'$) | 10 |
| Learning Ratio ($\alpha$) | 0.1 |
| Search limitation ($k$) | 10 |
| Initial spreading level ($H_{min}$) | 0.2 |
| Spread width ($H_{inc}$) | 0.05 |
| Evaluation Score | AIC |

We perform an evaluation of PBIL-RS in comparison with existing methods. We compare the performance of PBIL-RS with K2 that order restriction is evolved by genetic algorithms (K2-GA) [3], PBIL without mutations, and PBIL with three different mutation operators BM, TM, and PM. Parameter values used in the evaluation are shown in Table 1. Note that the mutation probability for BM, TM, and PM that performs the best is different for each mutation operators. Thus, we carefully chose those through preliminary experiments. For BM we use 0.005 that is the best performance mutation probability in range [0.001:0.2]. Similarly, for TM and PM, we use 0.1 and 0.002 that are the best in range [0.001:0.2] and [0.001:0.009], respectively.

Table 2 shows the comparison result summarizing the value of AIC calculated by each method. In Table 2, we show the performance of each method running 500 generations for two Bayesian Network models. In this result, we use the mean

Table 2: AIC Values at 500 Generations

| Methods | Bayesian Network Models | |
|---|---|---|
| | Alarm (37 events) | Pathfinder (135 events) |
| PBIL-RS | 8536.4 | 30138.6 |
| PBIL | 8627.2 | 30243.7 |
| PBIL + BM | 8563.1 | 35240.9 |
| PBIL + TM | 8654.3 | 34784.2 |
| PBIL + PM | 8582.9 | 33003.0 |
| K2-GA | 13347.7 | - |

Table 3: Computation Time (sec) for 500 Generations

| Methods | Bayesian Network Models | |
|---|---|---|
| | Alarm (37 events) | Pathfinder (135 events) |
| PBIL-RS | 8914.9 | 125195.6 |
| PBIL | 2791.9 | 75618.1 |
| PBIL + BM | 7992.4 | 156114.0 |
| PBIL + TM | 7920.7 | 161241.2 |
| PBIL + PM | 8724.9 | 142665.1 |
| K2-GA | 474327.0 | - |

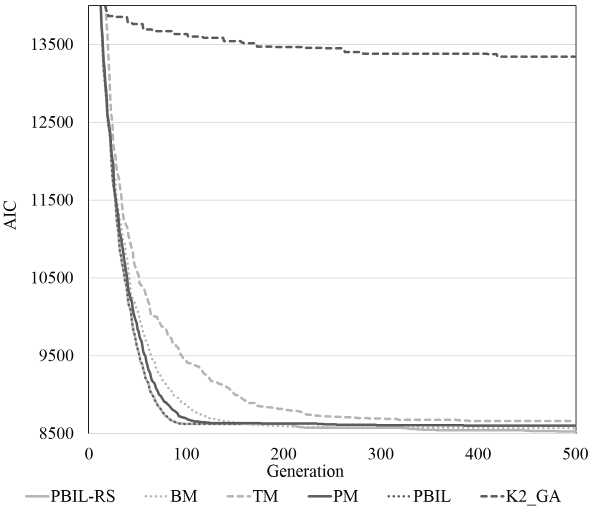



Figure 7: AIC Transition in Case of Alarm Network

Table 4: Variances of AIC Values at 400 Generations

| Methods | Bayesian Network Models | |
|---|---|---|
| | Alarm (37 events) | Pathfinder (135 events) |
| PBIL-RS | 508.3 | 43594.1 |
| PBIL + BM | 2708.1 | 56006.6 |
| PBIL + TM | 7458.2 | 235816.8 |
| PBIL + PM | 2498.9 | 108321.9 |

of 10 repetitions. Also, in Fig. 7, we show the transition of AIC values in the case of Alarm Network. From these results, we found that the PBIL-series methods perform far better than the traditional K2 although its order restriction is evolved by genetic algorithms, which proves the excellent ability of PBIL-based algorithms. Note that we could not compute the score of K2-GA for Pathfinder because it requires very large amount of time; it took 650 hours to proceed only 45 generations, whereas PBIL-RS took only 3 hours.

We also found that PBIL-RS has the best performance among those PBIL-based algorithms. This is because PBIL cannot continue searching after convergence (e.g., it finishes running at 160 generations in Alarm and 302 generations in Pathfinder), while BM, TM, and PM frequently change the searching area before exploring there deeply.

In Table 3, we show the execution time of those algorithms for 500 generations. PBIL finished in especially short time, which is because it finishes execution whenever the search space converges. On the other hand, K2-GA takes very long time because of large searching time of K2 algorithm. Except for those two, we found that the execution time of the variations of PBIL are comparable.

Table 4 shows the variance of AIC scores at the 400th generation with 10 repeated executions. From this result, we found that the variance of PBIL-RS is the smallest in both Alarm and Pathfinder, meaning that PBIL-RS most stably computes good solutions.

As above, we showed that PBIL-RS has the best performance among the methods compared in this evaluation. In the following sections, we examine the effect of several essential parameters on the performance of PBIL-RS, and investigate the key behavior of PBIL-RS that contributes to the superior ability.

## 4.3 Effect of Learning Ratio $\alpha$

We examine the effect of two essential parameters through several evaluations. First, we focus on the effect of Learning Ratio $\alpha$. We execute PBIL-RS with several learning ratios in range [0.05:0.7], and compare AIC values of Bayesian Network models.

Figure 8 shows the transition of AIC scores of each Bayesian Network models as generation proceeds. In Fig. 8, we show the average AIC values of 30 repetitions, where the horizontal axis represents generations, and the vertical axis represents the best AIC score found as generation proceeds. Also, in Fig. 9, we show the AIC scores at 500th generation for each learning ratios.

From those results, we found the property that the final AIC scores are better when learning ratio $\alpha$ takes lower values. Simultaneously, however, if $\alpha$ takes lower values, the speed to find better solutions goes slower than the case of higher $\alpha$. From this trade-off, we found that lower learning ratio $\alpha$ is better, but if we have limitation on the executable generations, we have to determine $\alpha$ carefully.

## 4.4 Effect of Spreading Level $H$

PBIL-RS increases the spreading level step-by-step to search larger search spaces to avoid falling into the same local mini-

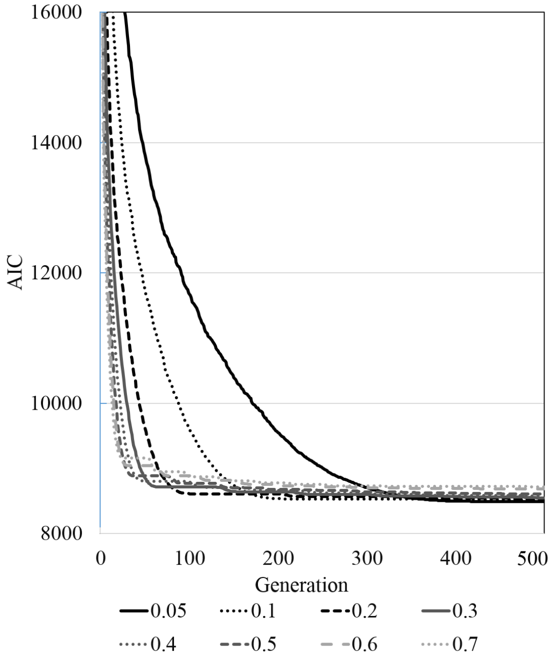



Figure 8: AIC Values under Variation of Learning Ratio $\alpha$

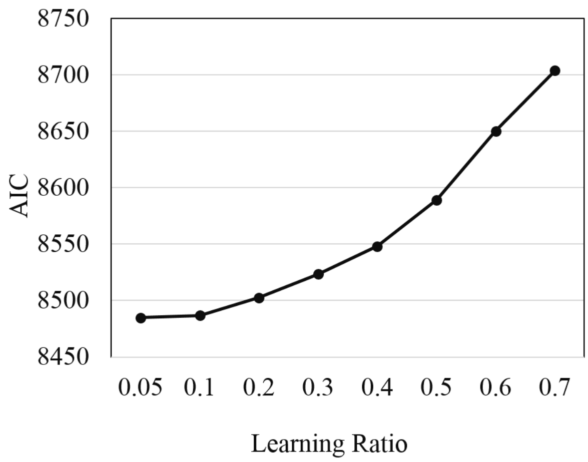



Figure 9: AIC Values with Learning Ratio at 500 Generation

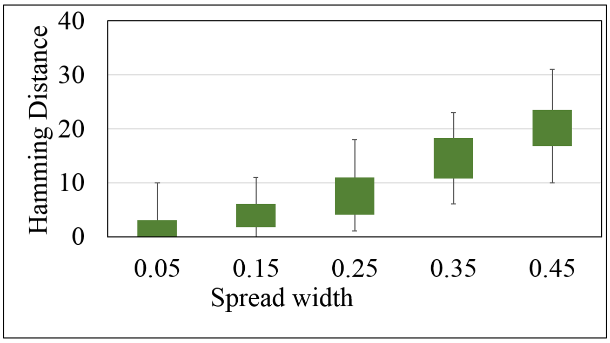



Figure 10: Hamming Distances among Convergence Points under Variation of Spreading Levels

mum area. In this section, we clarify the relationship between spreading level and the similarity of the Bayesian Network model structures in PBIL-RS. Also, by focusing on the convergence point of probability vector $P$, we find the key property of PBIL that show the superior performance of PBIL-RS.

Recall that each element $p_{ij}$ of the probability vector gets closer to 0 or 1 as the algorithm proceeds and generation grows. In most cases, $p_{ij}$ actually converges to 0 or 1, where $P$ always generate the same individual. Thus, we can regard each convergence point as a string of binary digits. We measure the similarity of two convergence points using Hamming distance of the corresponding strings in order to grasp the distribution of the convergence points under various values of the spreading level.

We examine the distance among convergence points under variation of spreading levels. Specifically, we varied the *initial spreading level* $H_{min}$ in [0.05:0.45] under fixed *spread width* $H_{inc} = 0$, and ran PBIL-RS with 1000 generations to examine the Hamming distance between every pair of the convergence points in a single run. Figure 10 shows the result as a box plot where the above and the below of the boxes show the maximum and minimum values, and the top and the bottom of the boxes show the values of the first and third quartiles of all the data. From Fig. 10, we see that the spreading level is clearly related to the Hamming distance between the convergence points, meaning that large spreading levels has an ability to change search space and to avoid converging to local minimum areas.

Next, we examine the behavior of the algorithm, especially on the timing at which good models are found. We ran PBIL-RS for 1000 generations with parameter values $H_{min} = 0.2$ and $H_{inc} = 0$. Figure 11 shows the result where the line represents the transition of *convergence level* $S$ and the dotted points represents the timing where the best model is updated. From Fig. 11, we see the basic behavior of PBIL-RS such that each time $P$ converges it spreads the search space. Please pay attention to the behavior of PBIL-RS that the point it updates the best models is concentrated on where the *convergence level* $S$ is very small. Figure 12 is the scatter diagram that shows the distribution of the updated points. Each dotted point represents the updated point, and the horizontal axis shows the the updated amount of AIC, i.e., the difference between the new AIC value and the previous one. From Fig. 12, one sees that the best Bayesian Network model is updated when *convergence level* $S$ is in the range of small values [0.0005:0.015]. The above results have clarified that deep exploration of local areas is an efficient strategy to find good Bayesian Network models.

## 4.5 Behavior in Varied Spread Level

In the previous section, we found two important properties of PBIL-RS that, first, large spread levels have an ability to avoid local minimum areas, and second, deep exploration of a local area is a preferable strategy to find good models, which support the superior performance of PBIL-RS.

In this section, we examine the behavior of the original PBIL-RS in which the value of spreading width is varied. We ran PBIL-RS in 3000 generations. Figure 13 shows the result

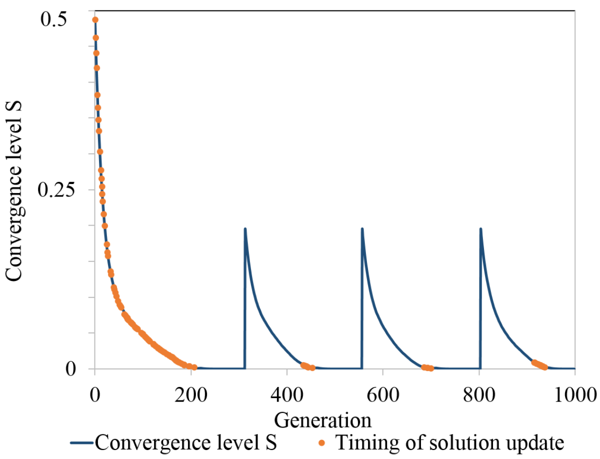



Figure 11: Transition of Convergence Levels, and the Updated Timing

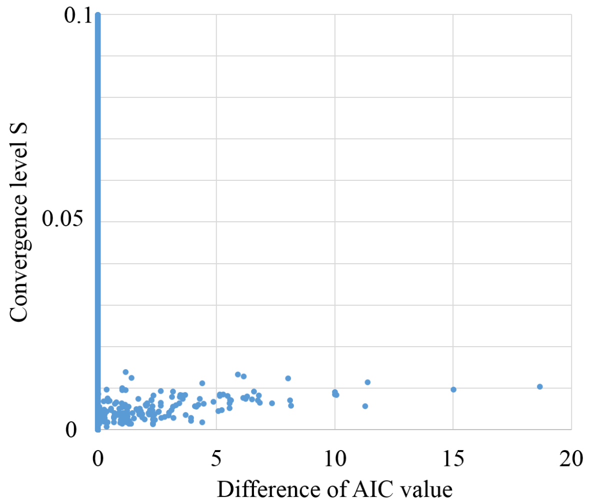



Figure 12: Convergence Levels at Each Update Point

where the line shows the transition of convergence levels as generation proceeds, and the dotted points shows the timing at which the best model is updated. From Fig. 13, we see that, as we found in the previous section, PBIL-RS finds the best model when the spread level comes to be low, and PBIL-RS continues finding better models even around 3000th generations by changing the exploring areas adaptively. We conclude that the good properties found in the previous section also work in the original behavior of PBIL-RS in which the convergence level is changed adaptively.

## 5 CONCLUSION

In this paper, we proposed a new algorithm called PBIL-RS, which is an algorithm to learn Bayesian Network models. PBIL-RS is an extension of PBIL that avoids convergence to local minimum solutions by means of spreading the search space repeatedly whenever it converges to a small area. Note that PBIL-RS is somewhat similar to the simulated annealing, but PBIL-RS is different from it in that PBIL-RS uses statistic property to find good solutions whereas the simulated annealing uses only the random effect. We evaluated the performance of PBIL-RS in comparison with existing algorithms, and we showed that PBIL-RS outperforms other existing algorithms regardless of the number of nodes in the Bayesian Network models used in the evaluation. In addition, we showed that the learning ratio significantly effects on efficiency of the algorithm, which clarified that selecting suitable learning ratio according to the planned execution time is important. Moreover, by examining the behavior of PBIL-RS, we verified that it properly controls the search space depending on the situation, and which leads to the superior performance.

As future work, more extensive evaluation using various Bayesian Network models is important. Especially, we would like to apply the models that include several thousands of nodes.

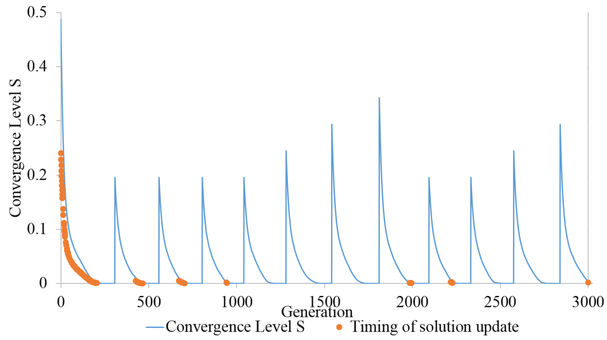



Figure 13: Behavior of PBIL-RS in Changing Convergence Levels

## ACKNOWLEDGEMENT

This work was partly supported by “the Program for Promotion of Stockbreeding” of JRA (Japan Racing Association).

## REFERENCES

[1] D.M. Chickering, D. Heckerman, C. Meek, “Large-Sample Learning of Bayesian Networks is NP-Hard,” Journal of Machine Learning Research, Vol.5, pp.1287-1330 (2004).

[2] G.F. Cooper, and E. Herskovits, “A Bayesian Method for the Induction of Probabilistic Networks from Data,” Machine Learning, Vol.9, pp.309-347 (1992).

[3] W.H. Hsu, H. Guo, B.B. Perry, and J.A. Stilson, “A Permutation Genetic Algorithm for Variable Ordering in Learning Bayesian Networks from Data,” In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO’02), pp.383-390 (2002).

[4] O. Barriére, E. Lutton, P.H. Wuillemin, “Bayesian Network Structure Learning Using Cooperative Coevolution,” In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO’09), pp.755-762 (2009).

[5] A.P. Tonda, E. Lutton, R. Reuillon, G. Squillero, and P.H. Wuillemin, “Bayesian Network Structure Learning from Limited Datasets through Graph Evolution,” In Proceedings of the 15th European conference on Genetic Programming (EuroGP ’12), pp.254-265 (2012).
[6] H. Handa, “Estimation of Distribution Algorithms with Mutation,” Lecture Notes in Computer Science, Vol.3448, pp.112-121 (2005).
[7] D.W. Kim, S. Ko, and B.Y. Kang, “Structure Learning of Bayesian Networks by Estimation of Distribution Algorithms with Transpose Mutation,” Journal of Applied Research and Technology, Vol.11, pp.586-596 (2013).
[8] S. Fukuda, Y. Yamanaka, and T. Yoshihiro, “A Probability-based Evolutionary Algorithm with Mutations to Learn Bayesian Networks,” International Journal of Artificial Intelligence and Interactive Multimedia, Vol.3, No.1, pp.7-13 (2014).
[9] R. Blanco, I. Inza, P. Larrañaga, “Learning Bayesian Networks in the Space of Structures by Estimation of Distribution Algorithms,” International Journal of Intelligent Systems, Vol.18, pp.205-220 (2003).
[10] H. Akaike, “Information theory and an extension of the maximum likelihood principle,” Proceedings of the 2nd International Symposium on Information Theory, pp.267-281 (1973).
[11] I.A. Beinlich, H.J. Suermondt, R.M. Chavez, G.F. Cooper, “The ALARM Monitoring System: A Case Study with Two Probabilistic Inference Techniques for Belief Networks,” In Proc. of Second European Conference on Artificial Intelligence in Medicine, Vol. 38, pp.247-256 (1989).
[12] D.E. Heckerman, E.J. Horvitz, B.N. Nathwani, “Towards Normative Expert Systems: Part I - the Pathfinder Project,” Methods of Information in Medicine, pp. 90-105 (1992).
[13] S. Baluja, “Population-Based Incremental Learning: A method for Integrating Genetic Search Based Function Optimization and Competitive Learning,” Technical Report CMU-CS-94-163, Carnegie Mellon University (1994).

(Received September 22, 2015)

**Yuma Yamanaka** received the B.E. degree from Wakayama University in 2015. He is currently a Master-course student in Wakayama University.

**Takatoshi Fujiki** recieved his B.E. and M.E. degrees from Wakayama University in 2010 and 2012, respectively. He is currently a doctoral course student in Wakayama University. He is interested in data mining, machine learning, and bioinformatics. He is a student member of IPSJ.

**Sho Fukuda** received his B.E. and M.E. degrees from Wakayama University in 2012 and 2014, respectively. He is currently working with Intec Hankyu Hanshin Co.Ltd. He is interested in Data Analytics with large data sets.

**Takuya Yoshihiro** received his B.E., M.I. and Ph.D. degrees from Kyoto University in 1998, 2000 and 2003, respectively. He was an assistant professor in Wakayama University from 2003 to 2009. He has been an associate professor in Wakayama University from 2009. He is currently interested in the graph theory, dis- tributed algorithms, computer networks, medial applications, and bioinformatics, and so on. He is a member of IEEE, IEICE, and IPSJ.

# Visualisation and Avoidance of Uneven Road Surfaces for Wheelchair Users

Hiroshi Jogasaki[*], Shinichiro Mori[**], Yoshitaka Nakamura[*], and Osamu Takahashi[*]
[*]School of Systems Information Science, Future University Hakodate, Japan
[**] Fujitsu Laboratories Ltd.
G3113001@fun.ac.jp

***Abstract*** - The number of people aged 65 and above is increasing in Japan, and the ratio of wheelchair use of such people is higher than that of the younger generation. Hence, wheelchair use has been steadily increasing. Manual wheelchairs are most commonly used because they have more reasonable prices than electric-powered wheelchairs; however, they are also more restrictive. In addition, it is better if users can trace in a comfortable road surface. For this purpose, we can exploit the global positioning system (GPS) sensors, acceleration sensors and gyro sensors that have recently been embedded in smartphones. People can collect diverse data from smartphone sensors and share their knowledge through a network. This paper proposes a suggesting system to avoid tracing in an uncomfortable road surface by which wheelchair users can sense the road surface profile. This suggesting system senses uneven area of the roads by a three-axis acceleration sensor and a GPS sensor installed in a smartphone, which considerably reduces the cost. Because discomfort levels vary among users, we have developed a mapping solution which indicates a user's discomfort places by an interactive method on the smartphone. The system helps users mitigate uncomfortable reactions caused by passing through uneven roads in wheelchairs.

***Keywords***: probe information system, navigation system for wheelchairs, smartphone, wheelchairs, discomfort level

## 1 INTRODUCTION

Japan is entering a period of decreasing birth rate and ageing population. Consequently, wheelchair use is increasing and the wheelchair market has diversified. Electric-powered wheelchairs are currently evolving and gaining popularity. However, their cost renders them inaccessible to many people. Most schools, hospitals and department stores continue to use human-assisted models. These models provide on-site mobility; however, users who are unfamiliar with the site and are navigating a road for the first time may encounter problems. Therefore, we propose a navigation aid for wheelchair users, and evaluate its ability in a series of manual wheelchair experiments.

Miyazaki compared the social inaccessibility among walking persons, users of electric-powered wheelchairs and manual wheelchair users [1]. Contemporary lifestyle revolves around social activities such as working at the office, shopping, school lessons and receiving medical treatment at hospitals. Miyazaki's results are summarised in the comparison maps in his paper. These maps present the Tama Area outskirts of Tokyo. Miyazaki found that persons with impairments encounter various barriers to activities in the urban environment. In particular, users of manual wheelchairs have limited access to public facilities such as hospitals compared with walking persons and people using electric-powered wheelchairs.

Even when access itself is not problematic, travelling down the street may be uncomfortable for wheelchair users.

Able persons cannot imagine the impact of road conditions on wheelchair users. Even on flat surfaces, wheelchairs are sensitive to the texture of the road. Because they are used for directional steering, the front wheels of typical wheelchairs are tubeless with a small diameter (not larger than 7 inches) [2]. However, this design increases the sensitivity to the road surface conditions. A small-diameter solid tyre greatly senses the surface irregularities.

By visualising the discomfort level of the road surface, users can navigate their wheelchairs to avoid these courses.

The discomfort level of wheelchair vibrations relates to not only the surface conditions but also the speed of the wheelchair and the weight of the user. Steep and sudden slopes pose additional dangers to wheelchair users as the speed of the wheelchair becomes harder to control. When users decelerate while approaching a bad road condition, they can mitigate their discomfort level. In outdoor environments, GPS sensors on smartphones can detect the user's position on the road. If the sidewalk is sufficiently wide, the conditions can differ on the same road. For this purpose, the precision of the GPS sensors must reach 1 m.

As micro electric mechanical system (MEMS) technology becomes more widespread, more smartphones are being installed with sensor technologies. Since 2000, MEMS technology has been incorporated in automotive acceleration sensors which detect a sudden deceleration to enable airbags, gyro sensors which guide drivers along safe routes and reduce hand vibrations when pushing digital camera buttons, and many other systems. The field of mobile sensing has greatly advanced in recent years; many studies have focused on incorporating such sensors in cars, bikes, bicycles and wheelchairs. Moreover, the number of smartphone users is increasing dramatically, and acceleration and gyro sensors are increasingly used for logging long-term data. With these tools, users can gather information on their circumstances and share that information among interested parties. Using their smartphones, people can develop a cost-effective, convenient networking system based on the embedded sensor technology. This study addresses the installation of three-axis acceleration sensors and GPS sensors on the smartphone, which detect bad road conditions, which would increase wheelchair vibration.

Our method detects uneven road surfaces that convey unpleasant vibrations to wheelchair users. The individual discomfort levels of users are determined by an interactive input method. Based on this input, the system indicates the potential discomfort zones on the road.

## 2 RELATED WORK

To the best of our knowledge, there are at least four related papers on navigation systems for wheelchair users. One system uses acceleration sensors on a smartphone; the others retrieve the logs from commercial three-axis accelerometers.

### 2.1 Indoor and Outdoor Navigation System for Disabled Persons

In 2012, Nattapob et al. created a navigation system based on static information, such as stairways and precipitous slopes [3]. The smartphone obtains values from a direction sensor and communicates to a wheel speed sensor through Bluetooth. However, this method is not designed for sharp drops in road level (Fig. 1), sidewalks under construction (Fig. 2), sections of old pavements (Fig. 3), or similar conditions which increase the discomfort of wheelchair users. Therefore, Nattapob's system cannot detect the discomfort level of a road surface [3].

### 2.2 Unevenness Evaluation of Sidewalk Pavement by Vibrational Acceleration of Wheelchair

In 2004, Okamura and colleagues demonstrated that the surface of a paved sidewalk causes undesired vibration, which may physically harm the human body [4]. Moreover, the vibrational acceleration is proportional to the travel speed. They found that a surface similar to that illustrated in Fig. 4 can be traversed by wheelchair users for one hour per day without causing side-effects. An example of such a sidewalk is photographed in Fig. 5 [4] - [6].

Okamura's group reported the following four findings.

1) The dominant frequency of the acceleration response to wheelchair vibration of the wheelchair is a near-integer multiple of the space size between the joints.

2) The magnitude of the acceleration is proportional to the speed of the wheelchair and is most marked in the vertical direction.

3) Lighter persons experience more vibration than their heavier counterparts.

4) The measured vibrational acceleration indicates the discomfort level of the vibration.

### 2.3 Spatiotemporal Life-log Mining of Wheelchair Driving for Visualising Road Accessibility

In 2013, Iwasawa's group tested the ability of three-axis accelerometers to capture the surface road conditions [8]. They classified the road surface as rough and smooth and displayed the results in the comparison maps between actual and estimated statuses of the ground surface.

However, the accelerators used in this study were not smartphone-based. Moreover, the study objectives were to visualise the road accessibility, not determine the comfort level of the road surface [7] [8].

Although these authors theoretically evaluated the comfort level of wheelchair users by the VAL (vibration acceleration level), we consider that discomfort levels should be alleviated by human interaction, as they depend on the individual sensibilities of users.

Figure 1: Sharp drop in the road level

Figure 2: Sidewalk under construction

Figure 3: Old pavement on a sidewalk

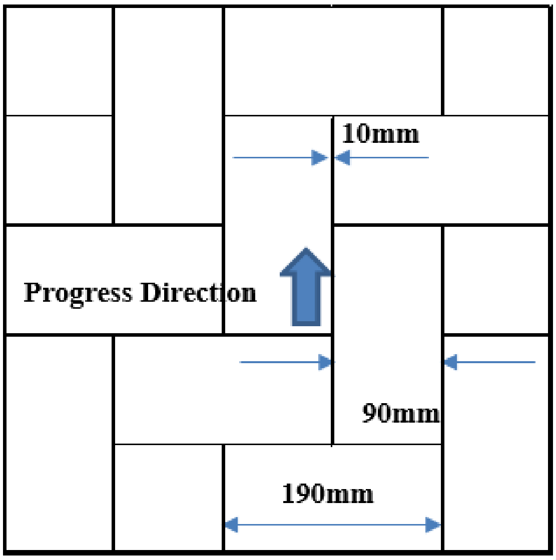



Figure 4: Uneven surface conditions on a paved sidewalk with tiles

Figure 5: Sidewalk paved with tiles

## 2.4 Determining the Discomfort Level from Inertia and Human Body Information of Wheelchair Users

In 2015, Isezaki and colleagues presumed the discomfort level from the body information of electric-powered wheelchair users [9]. They used a smartphone-based acceleration sensor and a biometric sensor which detects electric signals from the heart. However, they tested their system on an indoor walkway with and without a sharp drop, without varying the height of the drop. Their presumed discomfort level did not account for individual diversities.

Table 1: Advantages and disadvantages of related work

| | surface condition on the road | Sensor on the Smatrtphone |
|---|---|---|
| 2.1 Navigation | No | No |
| 2.2 Evaluation | Yes | No |
| 2.3 Visualization | Yes | No |
| 2.4 Presumption | Yes | Yes |

## 2.5 Summary of Related Work

Table 1 summarizes the advantages and disadvantages of the related work mentioned in this section.

Nattapob et al.'s method does not sense the road surface, which determines the discomfort level of the vibration. Okamura's group evaluated the vibrational damage, but their method requires accelerometers specific to this purpose. Iwasawa et al.'s approach is limited by a similar specificity; moreover, it targets the road rather than users' discomfort levels. Although Isezaki et al. adopted a smartphone-based approach, the acceleration and biometric sensors presume the discomfort level of electric-powered wheelchair users.

The method is unsuitable for manual wheelchair evaluation because manual wheelchairs introduce noise in the biometric sensor.

Given these limitations, we propose a low-cost method that precisely senses individual discomfort levels imparted by the road surface. The information is provided by smartphone sensors. The user determines the threshold acceleration beyond which he or she would feel uncomfortable while driving the wheelchair. The proposed method then proactively displays the potential discomfort sites on the road.

# 3 PROPOSED METHOD

This section details the study purpose and approach, and explains how our proposed system solves the problems in related papers.

## 3.1 Purpose and Approach

Problems with existing approaches include high introductory cost, and inability to detect individual differences in discomfort levels.

Our proposed system collects information at reasonable cost by a three-axis accelerometer and a GPS sensor installed on a smartphone. The discomfort imparted to the user by the road surface and the GPS position of the uneven surface is conveyed by an interactive method. Moreover, because the discomfort level is user-specified, the system assesses individual discomfort levels from the vibrational acceleration.

This study aims to collect the acceleration values on bumpy road surfaces and display the badly conditioned regions of the sidewalk on the map. The final target of this study is a navigation system by which wheelchair users can avoid uncomfortable routes in their future travels.

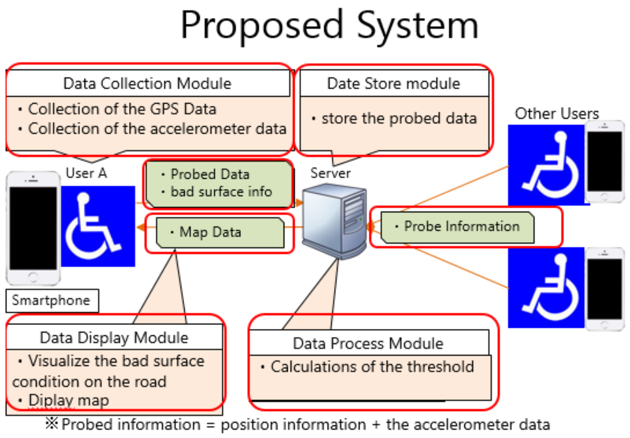



Figure 6: Modules of the proposed system

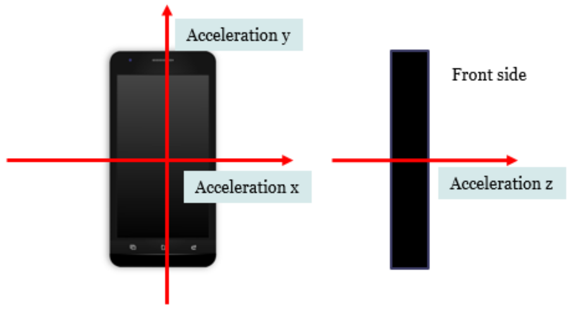



Figure 7: Acceleration components of the smartphone

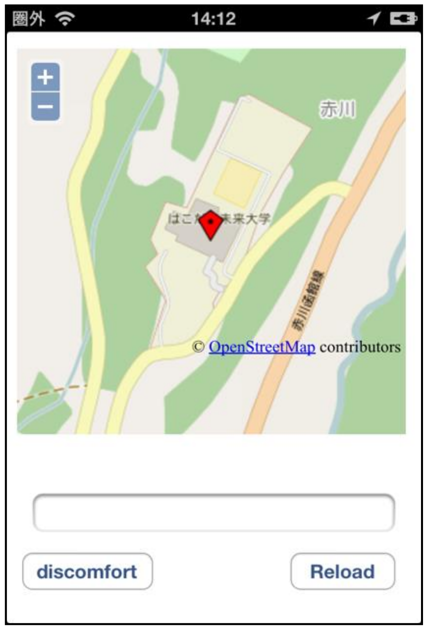



Figure 8: Data collection user interface on the smartphone

Table 2: Records sent to the server

| | |
|---|---|
| UserID | ID of each user |
| Latitude | Value from GPS |
| Longitude | Value from GPS |
| Acceleration Y | Value from Accelerometer |

## 3.2 Overview of the Proposed System

The proposed system comprises four components (Fig. 6). The interactive input method identifies a user’s individual discomfort level from historical data. The first module is the data collection module, which is installed on the smartphone. This module collects the sensed data from the three-axis accelerometers and the GPS sensors on the smartphone. The second module (the data processing module) operates on the server and calculates the discomfort threshold based on the gathered data. The third module is the data storage module, which receives data from the smartphone. This module also operates on the server. The fourth module (the data display module) presents the existing uneven road surfaces, on which the user’s discomfort level is likely to exceed the threshold. This information is presented on the smartphone map.

## 3.3 Visualising the Discomfort Site of the Road

In this subsection, we describe how our proposed method visualises the uneven road surfaces. The acceleration components of the smartphone are measured as indicated in Fig. 7. Once the unique ID is input to the system, the proposed system will store and calculate the threshold discomfort levels of the wheelchair user.

### 3.3.1 At the Start of Moving

The information is retrieved by an iPhone4 affixed to the left arm of the wheelchair. Data collection begins shortly after the user enters his or her unique ID into the smartphone screen (Fig. 8) and continues for 60 s. At 60-s intervals, the data are sent to the server, along with the records displayed in Table 2.

### 3.3.2 Data Collection

Each second, the smartphone acquires the latitude and longitude from the GPS sensor and the maximum and minimum accelerations in the $y$ direction from the three-axis accelerometer.

If the user experiences uncomfortable vibrations, he or she pushes the button on the screen of the smartphone. If the discomfort level is being tracked for the first time, that level will be stored in the database and assumed as the Min/Max threshold value (Fig. 9).

If the discomfort level has been previously tracked, it is stored after calculating the average value of the current and historical values, as shown in Fig. 10.

Table 3 lists the values that define a user’s discomfort profile. These values are stored in the database.

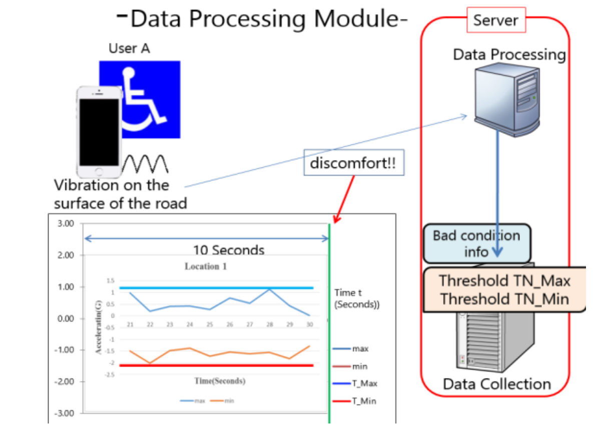



Figure 9: Data collection during first-time tracking

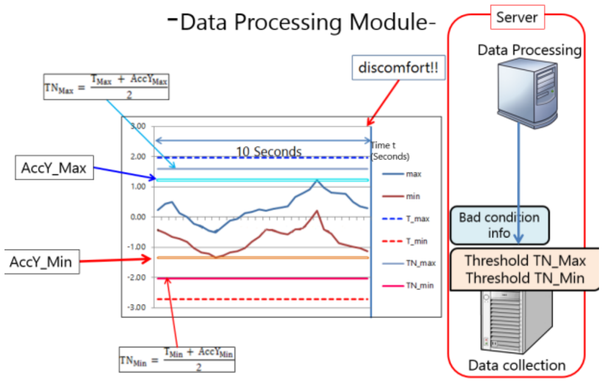



Figure 10: Data collection after multiple tracking

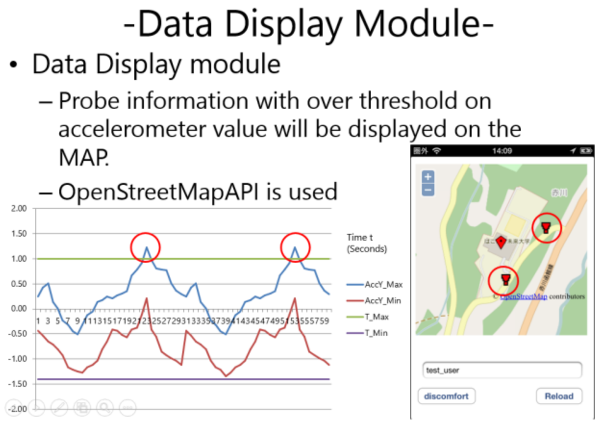



Figure 11: Data display on the smartphone screen


Figure 12: Experimental environment

Table 3: Records stored in the database

| | |
|---|---|
| UserID | ID of each user |
| Latitude | Value from GPS |
| Longitude | Value from GPS |
| Acceleration Y | Value from Accelerometer |
| Threshold max | Max threshold value of each user |
| Threshold min | Min threshold valu of each user |
| Comfort | information that feels discomfort |

### 3.3.3 Data Display

The system stores the values in the database and displays the information on the smartphone screen, as shown in Fig. 11. If the user's threshold is below that of multiple users in the same area, the map highlightsthe potential discomfort sites. The display API is OpenStreetMapAPI.

## 4 EXPERIMENTS AND DISCUSSION

In this section, we explain the experiments and simulation we conducted to confirm the effectiveness of the proposed method and discuss the results.

### 4.1 Experiments

We conducted two experiments. The first experiment was conducted inside the room and the second was conducted outside the building.

### 4.1.1 Experiment 1

Before creating the proposed method, we investigated the relationship between wheelchair vibrations and the road surface conditions. We found that the front tube-less tyre is sensitive to the surface unevenness. To assess the accuracy of sensing the road surface profile from the uncomfortable vibration level of the wheelchair users, we conducted the following tests.

(1) The smartphone was affixed to the left arm of the wheelchair. Data were collected by a HASC Tool [10] as shown in Fig. 12. Kawaguchi et al. organized a consortium-called 'HASC: Human Activity Sensing Consortium', and started a collaborative project for gathering a large scale human activity corpus. HASC Tool is a corpus tool, and its basic function was used. A schematic of the testing board, showing the varying spacing between several bars, is presented in Fig. 13. Each bar was 2 or 4 mm high. This experiment was called Experiment 1.

(2) We employed a common wheelchair with a 7-inch front tyre and a 24-inch rear tyre.

(3) We trialled 4 different patterns of the testing board.

(4) Each pattern was tested 5 times by 5 persons (yielding 200 trials: 4 patterns × 2 heights × 5 times × 5 persons).

(5) The test subjects were five students with different weights (55 kg, 58 kg, 64 kg, 65 kg and 73 kg).

In Experiment 1, the smartphone fixed on the left arm of the wheelchair detected the vibrations from the uneven testing board in the indoor environment. The testing board was designed to be detected from the ground by the front tubeless tyre of the wheelchair. The aim was to reveal the extent to which the smartphone's accelerometer can detect the surface profile from the vertical vibrational acceleration *y*. As revealed in the initial raw data, the surface signal was not precisely determined in this test (Fig. 14).

Noises in the accelerometer readings could be filtered by four periods of moving average. Figure 15 shows the acceleration profile after the filtering. Many of the noises in Fig. 14 were successfully eliminated by the filtering. A simplified graph is presented in Fig. 16. In almost all of the 200 trials, the exact space-size differences between the bars along the test board were derived from the acceleration sensor of the smartphone.

### 4.1.2 Experiment 2

After performing experiment 1, we gathered simulation data and verified the effectiveness of the proposed method. To sense the wheelchair's vibration from the real surface of the road, outdoor tests (experiment 2) were conducted from the front door of the university to the bus stop (Fig. 17). The five subjects made a round trip.

In Experiment 2, testers navigated the wheelchair along the road outside the university. Wheelchairs equipped with smartphones collected the acceleration and GPS values from the entrance to the bus-stop, along the route shown in Fig. 17. To identify whether the system could identify actual environmental features such as hearing the user's voice and perceiving the surface status of the road (smooth or rough), the test team followed the wheelchair users while taking a movie on a different smartphone.

The effectiveness of the method was evaluated in a simulation study based on Experiment 2.

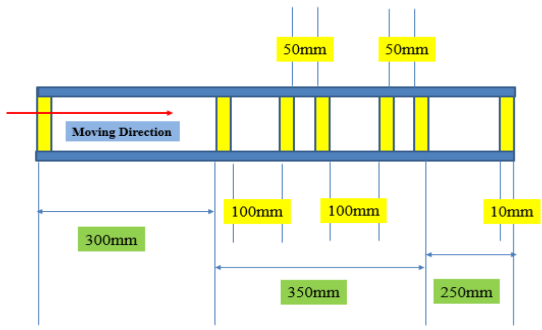



Figure 13: Schematic of the testing board with an uneven surface

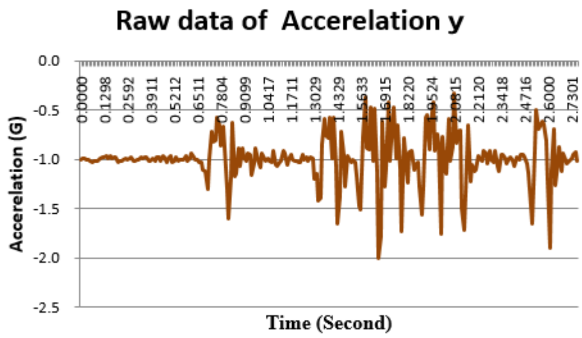



Figure 14: Raw data of vertical acceleration *y*

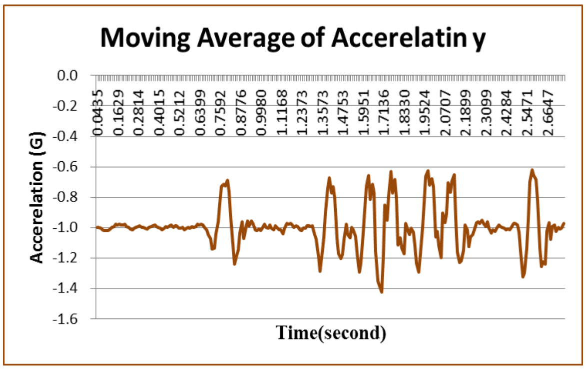



Figure 15: Moving average of vertical acceleration *y*

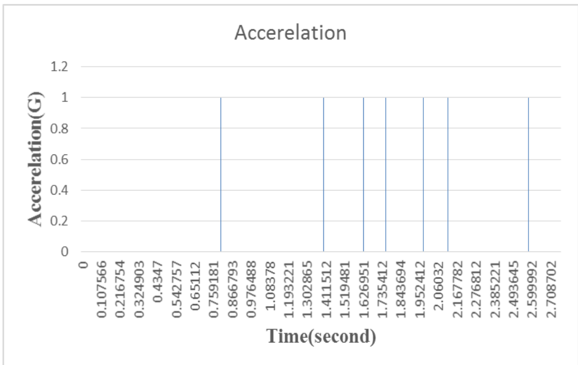



Figure 16: Simplified graph of vertical acceleration *y*


Figure 17: Experimental route from university entrance to bus stop

### 4.1.3 Results of Interview for Experiments

After conducting experiments 1 and 2, we surveyed each subject to evaluate the relationship between wheelchair vibrations and each individual wheelchair user's discomfort level.

Table 4 displays the format of a survey administered to the 5 testers. The testers rated three categories 'comfort', 'restfulness' and 'stability' on one of seven levels (extremely negative, moderately negative, slightly negative, neutral, slightly positive, moderately positive and extremely positive).

The responses were analysed by principal component analysis (PCA) [11]. PCA is a useful statistical technique that has found application in fields such as face recognition and image compression, and it is a common technique for determining patterns in high-dimension data. In the first analysis, we integrated comfort and restfulness into a measure of comfortability when using the wheelchair. In the second analysis, we integrated restfulness and stability into a measure of user safety. A scatter diagram of the analysis is plotted in Fig. 18. The levels of comfort and safety differed among the testers, although all testers moved along the same conditional testing boards and outdoor route.

## 4.2 Simulation Results

We performed simulation using data from experiment 2 to verify the effectiveness of the proposed method. We acquired logging data from the university entrance to the bus stop on the acceleration and GPS sensors embedded into the smartphone.

The wheelchair users encountered uneven sites at 27 s (no.1 in Fig. 19), 45 s (no.2 in Fig. 19) and 57 s (no.3 in Fig. 19) from the start of the route (university entrance). We then simulated this situation and determined the locations of the bumpy sites from the simulation results.

The tracking movie captured on the smartphone easily distinguished the regions of the bumpy sidewalk. An example is shown in Fig. 20.

From the wheelchair log data, we constructed the acceleration–time profile and identified the minimum and maximum accelerations.

The profiles around the first, second and third uneven sites are presented as Fig. 21, Fig. 22 and Fig. 23, respectively. The identified maximum and minimum values were selected as the threshold discomfort levels of the wheelchair's vibration.

Table 5 shows the *y* acceleration values stored by the smartphone. The discomfort level of the road surface was displayed on the smartphone screen.

## 4.3 Discussion of Experimental and Simulation Results

This subsection discusses the experimental and simulation results. The two experiments yielded the following findings. In experiment 1, in 200 trials, the signals of the testing boards were detected with an almost 100% success rate.

Table 4: Format of the interview

| Please mark your responses on the survey | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1. discomfort | extremely negative | moderately negative | slightly negativ | neutral | slightly positive | moderately positive | extremely positive | 1.comfort |
| 2. restless | extremely negative | moderately negative | slightly negativ | neutral | slightly positive | moderately positive | extremely positive | 2.restful |
| 3. unstable | extremely negative | moderately negative | slightly negative | neutral | slightly positive | moderately positive | extremely positive | 3.stable |

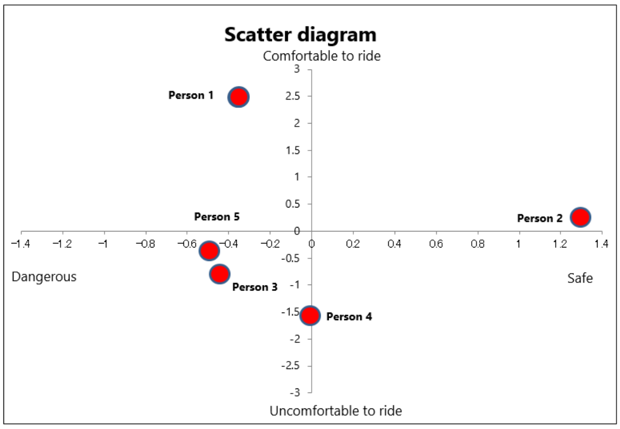



Figure 18: Scatter diagram of the principal component analysis

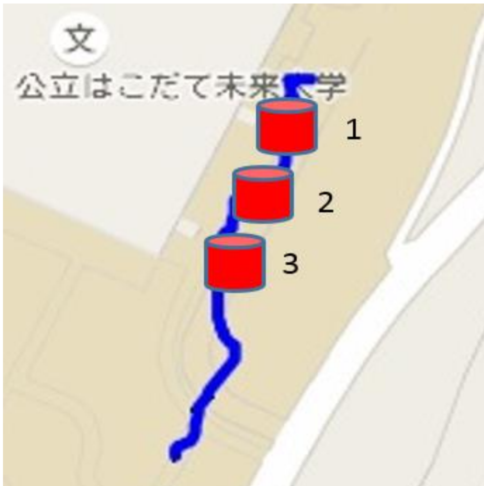



Figure 19: Experimental route from university entrance to bus stop, showing uneven sites

Figure 20: Example of a bumpy sidewalk

Hence, we could see that the wheelchair's vibration has a strong association with the surface profile of the road.

In experiment 2, we obtained accurate vibration and GPS data. This shows the effectiveness of the map as the output of the data displayed the potential discomfort location. The


Figure 21: MAX/MIN accelerations at the first uneven site along the outdoor route

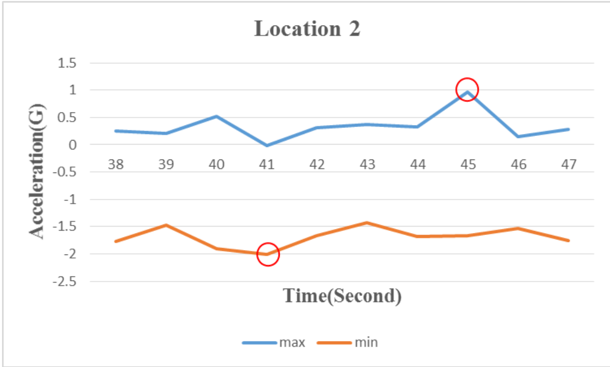



Figure 22: MAX/MIN accelerations at the second uneven site along the outdoor route

scatter diagram of interviews and tracking movies of experiment 2 showed that the comfort levels of wheelchair users are highly individual specific. Therefore, the vibration level deemed as uncomfortable cannot be standardised. Our simulation with the experiment 2 data showed the effectiveness of visualizing the potential discomfort levels for wheelchair users on the smartphone map using a sensor embedded into the smartphone.

Compared with existing methods, our method has several advantages. Existing studies of wheelchair navigation are beset by several problems. In particular, both the acceleration and GPS data must be obtained at reasonable costs, and the discomfort level derived from acceleration values is highly individualistic. Our proposed method addresses these issues.

However, there are other factors that have an influence on the discomfort level of wheelchairs are undetectable by acceleration and GPS data. One of the factors is sharp slopes. Without vibration, sharp slopes could be a hindrance to wheelchair users. Note that the current work does not address this issue.

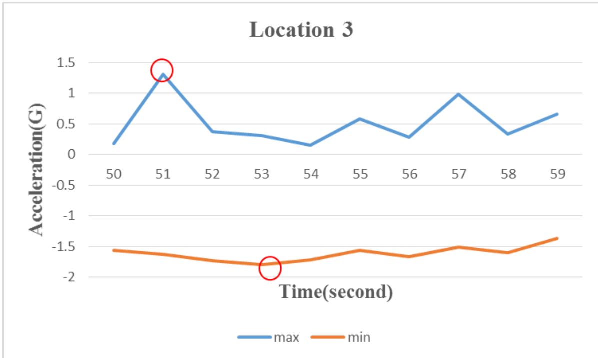



Figure 23: MAX/MIN accelerations at the third uneven site along the outdoor route

Table 5: Stored MIN/MAX accelerations reported as the thresholds

| | Real Value | | Threshold (after calcultion) | |
|---|---|---|---|---|
| | Acc Y min (G) | Acc Y max (G) | Acc Y min (G) | Acc Y max (G) |
| Location1 | -2.0084475 | 1.14548875 | -2.0084475 | 1.14548875 |
| Location2 | -2.007637 | 0.96374525 | -2.00804225 | 1.054617 |
| Location3 | -1.79909875 | 1.038414 | -1.9035705 | 1.0465155 |

## 5 CONCLUSIONS

In this study, we proposed a method for visualization and avoidance of uneven road surfaces for manual wheelchair users using smartphones with an embedded acceleration sensor and a GPS sensor. We implemented the system of the proposed method and conducted experiments and simulation to confirm its effectiveness.

We proposed a method displaying potentially uncomfortable zones as surface information on the road map of the smartphone. Because every trigger is initiated by the wheelchair user, this system will provide individual assistance.

In future work, the system must be evaluated on a real sidewalk. We will also create a navigation system to evaluate the effectiveness of the method.

At present, our method relies on the wheelchair users' own activities. Ideally, we should find a proactive measure, by which probe-wheelchairs could check the road status before the road is accessed by many users. For a more realistic navigation, we also need to identify the slope by gyro sensors and consider the scalability of the servers and collectiveness of the probed data.

## ACKNOWLEDGEMENTS

This work's experiments were partly supported by students headed by Yuta Ibuchi, a student of the School of Systems Information Science, Future University Hakodate.

The authors would like to thank Enago (www.enago.jp) for the English language review.

## REFERENCES

[1] H. Miyazaki, "The Urban Built Environment and Inaccessibility: A Case Study of Early Developed Area of Tama New Town, Tokyo," The Human Geographical Society of Japan, Vol 56, No. 1, pp.1-20 (2004) (in Japanese).

[2] Spinal outreach team and school of health and rehabilitation sciences university of Queensland, Manual

wheelchairs–Information resource for service providers (2014).
[3] W. Nattapob and Wakahara Toshihiko, "Indoor and outdoor navigation system for disabled," IPSJ SIG Technical Report (2012) (in Japanese).
[4] M. Okamura and N. Fukada, "Study on unevenness evaluation of sidewalk pavement with vibration acceleration of wheel chair," Japan Society of Civil Engineering, Pavement Engineering, Vol 9, pp17-24 (2004) (in Japanese).
[5] M. Okamura, "Effect of joint of tile pavement on vibration of wheelchair, Japan," Society of Civil Engineering, Vol 64, No.1, pp.237-246 (2008) (in Japanese).
[6] M. Okamura, "Study on performance indexes of sidewalk pavement by focusing on ride comfort of wheelchair," Japan Society of Civil Engineering, Pavement Engineering, Vol 14, pp.189-194 (2009) (in Japanese).
[7] Y. Iwasawa and I. E. Yairi, "Life-Logging of Wheelchair Driving on Web Maps for Visualizing Potential Accidents and Incidents," The 26th Annual Conference of the Japanese Society for Artificial Intelligence, 3D2-R-13-9 (2012) (in Japanese).
[8] Y. Iwasawa and I. E. Yairi, "Spatiotemporal Life-Log Mining of Wheelchair Driving for Visualization Accessibility of Roads," The 27th Annual Conference of the Japanese Society for Artificial Intelligence, 1D3-5 (2013) (in Japanese).
[9] T. Isezaki, A. Miyata, T. Watanabe and T. Tanaka, "Presumption on discomfort level based on the inertia and human body information of the users of the wheelchairs," IPSJ SIG Technical Report Vol 2015-GN-95 No.18 (2015) (in Japanese).
[10] N. Kawaguchi, N. Ogawa, Y. Iwasaki, K. Kaji, T. Terada, K. Murao, S. Inoue, Y. Kawahara, Y. Sumi and N. Nishio, "HASC challenge: Gathering large scale human activity corpus for the real-world activity understandings," Proc of ACM AH (2011), pp.27:1-27:5 (2011).
[11] L.I. Smith, A tutorial on principal components analysis, Student tutorials on Computer Science Course at Otago University (2002) http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf

(Received September 30, 2015)
(Revised November 26, 2015)

**Hiroshi Jogasaki** received his B.E. in Computer Engineering from Waseda University in 1987. He entered IBM Japan in 1987 and is currently working at the UK company of Harvey Nash Japan KK as Business Development Director. He is working towards a Ph.D. at the School of Systems Information Science, Future University Hakodate. He is a member of IPSJ.

**Shinichiro Mori** received his B.E. from Kansai University in 1987. He received a Ph.D from Shizuoka University in 2011. He entered Fujitsu Limited in 1987 and transferred to Fujitsu Laboratories Ltd. in 2003. He has engaged in the research of positioning technology by way of the research of production robots and mobile phones/in-vehicle terminals. At present, he is the Senior Researcher.

**Yoshitaka Nakamura** received his B.E., M.S. and Ph.D. from Osaka University in 2002, 2004 and 2007, respectively. He is currently a research associate at the School of Systems Information Science, Future University Hakodate. He is a member of IEEE and IPSJ.

**Osamu Takahashi** received his M.S. from Hokkaido University in 1975. He worked for NTT research laboratory and NTTDocomo research laboratory. He is currently a professor at the Department of System Information Science at Future University Hakodate. His research interests include ad-hoc networks, network security and mobile computing. He is a fellow of IPSJ and member of IEEE and IEICE.

# A Lump-sum Update Method as Transaction in MongoDB

Tsukasa Kudo[†], Masahiko Ishino[‡], Kenji Saotome*, and Nobuhiro Kataoka**

[†]Faculty of Comprehensive Informatics, Shizuoka Institute of Science and Technology, Japan
[‡]Faculty of Information and Communications, Bunkyo University, Japan
* Hosei Business School of Innovation Management, Japan
** Interprise Laboratory, Japan
kudo@cs.sist.ac.jp

***Abstract*** - Along with the progress of the cloud computing, it became necessary to deal with various and large quantity data in the distributed database environments. So, the various NoSQL databases have been proposed and put to practical use. However, as for the NoSQL databases, since it supports the distributed environment, the integrity of the database update is basically guaranteed only by the object unit. Therefore, there are serious restrictions to update the plural objects as a transaction. On the other hand, it is often necessary to perform the lump-sum or long-time update as a transaction in business systems. In this paper, we propose a method to update plural objects in a lump-sum as a transaction in MongoDB, which is a kind of the NoSQL database. Furthermore, through the evaluations by a prototype, we confirmed that the lump-sum update can be executed as a transaction without the latency of the concurrent object unit update.

***Keywords***: database, NoSQL database, MongoDB, transaction processing, batch processing, concurrency control

## 1 INTRODUCTION

Nowadays, a large amount of data has been published, and it is utilized in various fields as big data. As a feature of big data, Volume (huge amount), Velocity (speed), Variety (wide diversity) have been pointed [8]. For example, large amounts of data, such as in the online shops and the video sharing sites, must be accessed efficiently by worldwide users, though it has more complex data structures than the conventional relational databases, including images and videos as well as texts.

To cope with this situation, various kinds of NoSQL (Not Only SQL) databases has been proposed and put to practical use [12]. As for the NoSQL database, to achieve the above-mentioned feature for the above problems, it is composed as the distributed database having a large number of servers. That is, it ensures the efficiency and reliability by redundancy such as replication and so on. Also, for example, MongoDB, which is a kind of the NoSQL database, is the document-oriented database and its structure is not defined by the schema. So, it is possible to add necessary attributes to its each data at any time and to manipulate various kinds of data flexibly [1].

On the other hand, unlike the relational database management system (hereinafter, "RDBMS"), it is not guaranteed to maintain the ACID property of the transaction processing in the case of the plural data manipulation. That is, it is generally maintained only on the individual update units called the atomic object. Also, as for the distributed environment, only the eventual consistency is guaranteed, that is, its consistency is not maintained until the completion of all the data manipulation including such as the synchronization of the replication [13]. They cause the serious restrictions on the data manipulation, for example, the intermediate result of the update having no consistency can be queried.

Here, even as for the RDBMS, there is a problem about maintaining the ACID properties in the case to update a large amount of data associated mutually in a lump-sum. That is, since the lock method is used to perform above-mentioned data manipulation concurrently with the other update, it causes the long latency of the latter update. For this problem, we have proposed the temporal update method using the transaction time database that manages the history of the time series of the data, and shown that it is possible to maintain the ACID properties without this long latency even in the above-mentioned case [6]. In this method, each update result of plural transactions is saved, and only the valid results are queried after the update completion. So, it is expected that we can update the plural data in MongoDB as the single transaction by the method of applying the temporal update method.

In this paper, firstly, we propose an update method for MongoDB, which utilizes the concept of the temporal update. And, our goal in this paper is to show that the efficient lump-sum update maintaining the ACID properties can be realized even in the above-mentioned case. In other words, by this method, we can update the plural data in a lump-sum as the single transaction, which was difficult to be executed by the conventional method of MongoDB.

The remainder of this paper is organized as follows. In Section 2, we show the problem of MongoDB about the concurrency control, and the abstract of the temporal update. In Section 3, we propose the lump-sum update method for MongoDB. In Section 4, we show the implementation and evaluations of this method, and show the considerations about this evaluation results in Section 5.

## 2 CONCURRENCY CONTROL OF MONGODB AND TEMPORAL UPDATE

### 2.1 Target Case of Data Update Process

Currently, many databases of mission-critical business systems are built by the RDBMS, and a lump-sum updates of a large amount of data spanning a long period of time are often performed. For example, in the banking systems, there is a large amount of account transfer business, which is entrusted by the card companies and so on. Meanwhile, users update the database immediately for their deposits and withdrawals by the ATM. These processes are not only performed simultaneously by many users, but also the users in the latter case are sensitive to the delay in the response of the system. Furthermore, at present, these processes are provided as the nonstop services. That is, the both of these processing have to be executed concurrently. However, since the lock method is generally used to maintain the ACID properties of the transactions, there is the problem that the users are often kept waiting for a long while by the update of large amount of data.

Here, the ACID properties are the properties that the transactions should maintain, and it is composed by the following four properties [2].

**Atomicity** Transactions execute all or nothing.

**Consistency** Transactions transform a correct state of the database into another correct state.

**Isolation** Transactions are isolated from one another.

**Durability** Once a transaction commits, its updates persist in the database, even if there is a subsequent system crash.

For the above-mentioned problem due to the lock method, the mini-batch is used widely to shorten the wait time of users, by which the process for a large amount of data update is divided into several update processes and they are performed one after another. So, the long wait does not occur.

Meanwhile, we showed that the ACID properties cannot be maintained by the mini-batch in the case where the data associated mutually. Furthermore, we proposed the temporal update method to update the data with maintaining the ACID properties even in this case [5]. On the other hand, as well as the mini-batch, the ACID properties of the transactions are maintained only on the update of the atomic object in MongoDB. So, in the case of updating plural data in a lump-sum, it has been pointed out that the same problems as the mini-batch occurs [1].

That is, the aim of this paper is to propose the update method for MongoDB, in which we apply the concept of the temporal update method, to update plural data in a lump-sum with maintaining the ACID property. In this paper, "online entry" represents the update on the atomic object, in which the ACID property is maintained in MongoDB; "batch update" represents the update composed of plural data manipulation in a lump-sum, in which the ACID property is not maintained to the contrary. Incidentally, as for the relational databases, the former corresponds to the update of the single transaction; the latter corresponds to the update composed of the plural transactions such as the mini-batch.

In this section, we show these related works below: first, the overview of MongoDB, and the issue intended by this paper; second, the temporal update method for the relational databases.

```
{ "_id" : 1, "name" : { "first" : "Tsukasa", "last" : "Kudo" },
"address" : "Hukuroi-shi, Shizuoka"}
```

Figure 1: Composition of MongoDB document.

## 2.2 Overview of MongoDB and Issue in Concurrency Control

MongoDB is a kind of document-oriented NoSQL database, which data is the documents expressed by JSON (JavaScript Object Notation) format shown in Fig. 1 [15]. The document is composed of the fields. For example, in this figure, {"_id": 1} is a field, of which identifier is "_id" and value is 1. Here, "_id" corresponds to the primary key of the relational database. And, the field is able to have a nested structure. For example, the name field (name) in the figure is composed of the following fields: the first name field (first) and the last name field (last). Since the document structure of MongoDB is not defined by schema, any necessary fields can be added to any document at any time. So, each document is able to have different fields except "_id". Furthermore, since it is possible to store the various kinds of objects, such as images and videos to its fields, it can handle a variety of data compared to the RDBMS. Here, the set of documents is called the collection. So, the collection and document correspond to the table and record in the relational database, though it is not strict.

As for the data manipulations, the following CRUD operations are provided as well as the RDBMS: insert, find (corresponding to select), update and remove (corresponding to delete). Furthermore, since findAndModify command is also provided to execute both of the query and update exclusively, they can be executed as the atomic operation. That is, the update of the atomic object such as the single document can be performed as the single transaction.

However, unlike SQL of the RDBMS, it does not provide the command to update the plural documents as the single transaction. That is, there is a problem that the ACID properties of the transaction, especially the isolation and atomicity, cannot be maintained in the case where the plural documents are updated in a lump-sum.

For this issue, two phase commit protocol is shown [10]. In this method, for example, in the case of performing the account transfer from the account $X$ to the account $Y$ of a bank, its processing ID is saved in the document of the account transfer management collection, which has the status about this processing: initial, pending, applied and done. These accounts are updated one after another in pending; meanwhile, this processing ID is saved to the documents of these accounts, and the updating accounts can be managed. Then, the status transit to applied. In the case of successful completion, this processing ID is deleted from the documents to exit. And, in the case of abnormal termination, the compensation transaction is performed to cancel the updates of the accounts and recover to a consistent state [10].

It is considered that this method is same as the saga for the RDBMS [9], which executes a mass update sequentially as divided plural update set. And, in the case of failure, the compensation transactions are executed to recover the data.

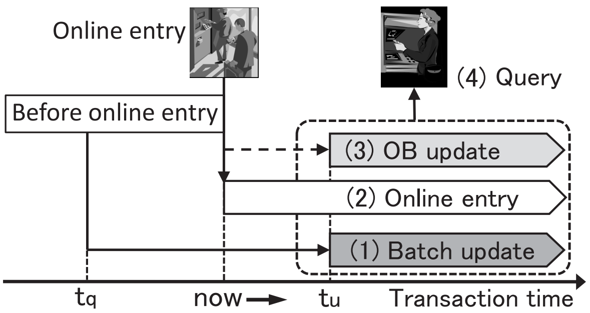



Figure 2: Data manipulation by temporal update.

On the other hand, it has been shown that the ACID properties of the transaction is not maintained in the case of the concurrent execution with the other transactions [3]. For example, in the case where the failure occurs in the transaction after the account $X$ was updated and the result was queried by the other transaction, the former transaction must be canceled. However, since the result of this transaction has been already queried by other transactions, it causes the problem in the actual system operations, such as the cascading aborts. This is due to the fact that the entire update cannot be processed as the single transaction in the same as the mini-batch.

## 2.3 Temporal Update Method for Relational Database

In Section 2.1, we mentioned that we proposed the temporal update method for the problem as for the mini-batch in the RDBMS. In Fig. 2, we show the data change of the time series about the transaction time in this method [7]. The concept of the temporal update method is the following: all the update results of each kind of transactions are saved, in which the online entries and batch update are included; and only the valid data is queried [5], [6].

This method utilizes the concept of the transaction time database, which is a kind of temporal database that manages the time history of its data [14]. And, the transaction time expresses when some fact existed in the database, so its relation is expressed by $R(K, T_a, T_d, A)$. Here, $K$ shows the primary key attribute of the data of above-mentioned fact; $T_a$ shows the transaction time when the data was inserted into the database; $T_d$ shows the transaction time when it was deleted from the database; $A$ shows the other attribute. In other words, though the data is deleted logically from the database by setting the deletion time to $T_d$, it remains physically in the database. So, all the CRUD operations on the data of the database can be managed as a history of the time series. Here, until the data is deleted, the value $now$ is set to $T_d$, which indicates the current time [16].

The feature of this method is that we avoid the conflicts between the batch updates and online entries by expanding the concept of the transaction time into the future. That is, as for the batch update, the data at the past time $t_q$ is queried, and the processing result is stored at the future time $t_u$. On the other hand, as for the online entries, the data at the current time $now$ is manipulated. Thus, the conflict between the both processing can be avoided without using the long locks.

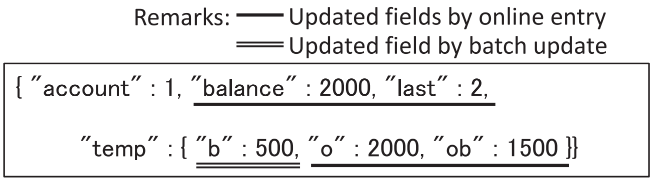



Figure 3: Data structure of document in proposal method.

Here, the batch update processing must be applied to the result of the online entries performed between the time $t_q$ and $t_u$. So, the batch update processing is applied individually to the results of the online entries, and these processing results are stored into the database as the OB update in Fig. 2.

As a result, three kinds of update result data is stored at the time $t_u$ as shown in Fig. 2: (1) the batch update, (2) the online entry and (3) the OB update. Therefore, the valid data has to be queried by the query processing, which is shown by (4) query. It is achieved by querying these data in the following order of priority: the OB update, online entry and batch update. To be concrete, in the case where both of the batch update and online entry are executed, the OB update result is queried; in the case of only the batch update, its result is queried; in the case of only the online entry, its result is queried; Therefore, we can query the same update result as in the case where the batch update is executed on the online entry results at the time $t_u$, without the long latency of the online entry by the lock method.

Moreover, to apply this method to the distributed database environment, we improved it not to have to determine the completion time $t_u$ beforehand [7]. It was implemented by the view, which has the feature that the above-mentioned valid data is changed at the batch update completion time $t_u$.

# 3 PROPOSAL OF LUMP-SUM UPDATE METHOD FOR MONGODB

In this section, we propose a lump-sum update method for MongoDB, which is based on the concept of the temporal update method mentioned in Section 2.3. First of all, since there is the restriction about the transaction processing in MongDB as mentioned in Section 2.2, we have to adopt the method to fit the characteristics of MongoDB.

Concretely, since the plural document cannot be updated as a transaction in MongoDB, the following modification is necessary. First, all the update result must be stored in the single document: the online entry, batch update and OB update. So, the data structure must be modified. Second, for the same reason, the transaction time database cannot be constructed. So, the update process must be managed according to the processing stage, instead of the transaction time.

## 3.1 Data Structure of Document in Proposed Method

If we applied the temporal update intended for the RDBMS to the lump-sum update in MongoDB, multiple documents of the update result would be created for one fact of the real world: by the online entry, batch update and OB update. It means, for example, in order to execute the OB update with the online entry, the two documents represented by (2) and (3) in Fig. 2 must be manipulated as one transaction. However, as mentioned above, plural documents cannot be updated as one transaction in MongoDB. So, if we applied this method to MongoDB just as for the RDBMS, the problem would occur: the consistency among data is not maintained. Meanwhile, since the document structure of MongoDB is not defined by the schema, its fields can be added flexibly. And, more importantly, the data manipulations on the different fields of the same document do not conflict.

For this reason, in this study, we propose the following lump-sum update method for MongoDB as shown in Fig. 3. In this method, all the results of the online entry, batch update and OB update are saved in the same document, and each field is indicated by "o", "b" and "OB". And, the valid field is queried in the same way as the temporal update for the RDBMS.

Here, as shown in Section 2.1, we defined the online entry as the process updating each document individually. And, it can be executed by the transaction feature of MongoDB. Furthermore, OB update, which accompanies with the online entry, updates the same data as the online entry. So, it can be executed by the transaction feature, though both of the OB update and online entry cannot be executed by the transaction feature in a lump-sum. On the contrary, since the batch update updates plural documents in a lump-sum, it cannot be executed by the transaction feature as the whole process.

For example, Fig. 3 shows the document of the balance of the deposit account: the account number (account), balance (balance) and update number (last). In addition, in the following, we omit to write "_id" of documents. Here, "last" corresponds to the time stamp, and it is increased by one for each update of the balance. In addition, it is used for the optimistic concurrency control as described later. Also, "temp" is the temporal field that is added temporarily during the execution of the batch update, and the update results are inserted to the corresponding field with the processing classification: the field of identifier "b" is for the batch update; "o" is for the online entry; "OB" is for the OB update. In addition, "balance" and "last" fields are also updated by the online entry.

As shown in Fig. 3 by the underline and double underline, the online entry, including the OB update, updates the different fields from the batch update field. So, there is no conflict between these update processing. Then, the valid update result can be queried similarly to the temporal update method, that is, by querying the update result data with the following priority: the OB update, online entry and batch update.

In summary, as for the temporal update method in RDBMS, we avoid the conflicts between the batch update and online entry by storing the each update results to the different records. In this method for MongoDB, we avoid this conflicts by storing each update result to the different fields.

## 3.2 Transition of Processing Stage in Proposed Method

In MongoDB, since plural documents cannot be updated as the single transaction, the transaction time database also

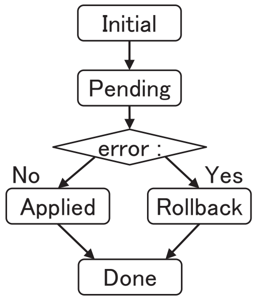



Figure 4: Processing stage transition in update

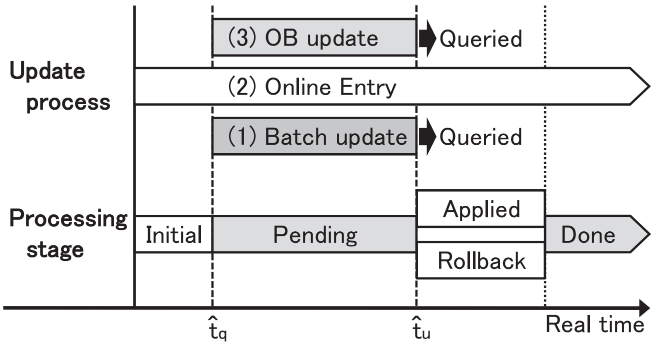



Figure 5: Correspondence between update process and processing stage

cannot be composed. For example, in Fig. 2, two data is manipulated to update one fact at time $now$: the data "Before online entry" is logically deleted, and the update result "(2) Online entry" is added. In other words, the transaction time database needs to manipulate two data as a single transaction.

On the other hand, as for the temporal update in RDBMS, as shown in Fig. 2, since the batch update starts at the time $t_q$ and completes at $t_u$, the OB update must be executed during this time period, which accompanies with the online entry. Furthermore, the batch update is performed to the data of the transaction time $t_q$. That is, these control is required for the proposed method, too.

To address this issue, we define the following four processing stage like the two phase commit protocol in MongoDB, which was shown in Section 2.2, as shown in Fig. 4. And, the update processing is performed with transitioning among them sequentially: "initial" shows that the stage is before batch update; "pending" shows it is during the batch update; "applied" shows batch update has completed, and the data of temporal field is being reflected to the regular field; "done" shows all the processing has completed. Incidentally, in the case where the failures occur in the batch update processing, the processing stage transitions from "pending" to "rollback". In the "rollback" stage, the batch update results are canceled, and the processing stage transitions to "done".

We show the correspondence between the update process and processing stage as for the "real" time in Fig. 5. Hereinafter, we use real time $\hat{t_q}$ corresponds to $t_q$ in Fig. 2, and $\hat{t_u}$ corresponds to $t_u$. At the transition time $\hat{t_q}$ from "initial" to "pending", the batch update and OB update start. And, at $\hat{t_q}$ the time from "pending" to "applied", the both complete.

```
{ "account" : 1, "balance" : 1000, "last" : 1}
```

(1) Data at end of initial

```
{ "account" : 1, "balance" : 2000, "last" : 2,
      "temp" : { "b" : 500, "o" : 2000, "ob" : 1500 }}
```

(2) Data at end of pending

```
{ "account" : 1, "balance" : 1500, "last" : 3}
```

(3) Data at end of applied

```
{ "account" : 1, "balance" : 2000, "last" : 2}
```

(4) Data at end of rollback

Figure 6: Data change in update

Here, as mentioned above, since MongoDB is not the transaction time database, the batch update cannot query the data history at the time $\hat{t_q}$. So, in the case where the target data is updated by the online entry before the batch update, this batch update must perform on this update result. However, the result of the OB update, which accompanies with the online entry, reflects both of the batch update and online entry results, and finally this is queried based on the query priority. So, the query result of the proposed method is same as the result of the temporal update method.

Here, the query data as of the proposed method is decided at the time $\hat{t_u}$, and the query results do not change after $\hat{t_u}$. That is, as shown in Fig. 5, the batch update and OB update result can be also queried corresponding on the query priority. The transition from "applied" and "rollback" to "done" means only the delete completion of the unnecessary intermediate results.

Figure 6 shows the data at the end of each processing stage. (1) shows the data at the end of "initial". Since it is prior to batch updates, "temp" field does not exist. Also, (2) shows "pending". Since the batch update has completed, the data has been set to temp field. In the case of this figure, the batch update debited 500 from the account. Meanwhile, the online entry deposited 1000 to the same account, and OB update debited 500 from this result. Then, all the results were stored in the temp field. At this time, "balance" and "last" fields have been also updated by the online entry. Incidentally, since the value is set only to the fields corresponding to the executed updates, all the fields of temp field are not always set.

While the processing stage is "applied", the valid data is queried by the online entry transactions. In the case of this figure, the balance of 1500 in "ob" field is queried. Furthermore, in this stage, the valid data is reflected into balance field, then temp field is deleted. This is the processing for the next batch update. At the end of this stage, each field has the value shown in (3), and 1500 is set to balance field, which is the result of the OB update. In this way, the query results of the online entry do not change through this stage.

On the other hand, in the case where the batch update processing fails, the processing stage transitions to "rollback". In this stage, only balance field is queried by the online entry continually; temp field is ignored. And, temp field is deleted without affecting the online entry. So, when the rollback has

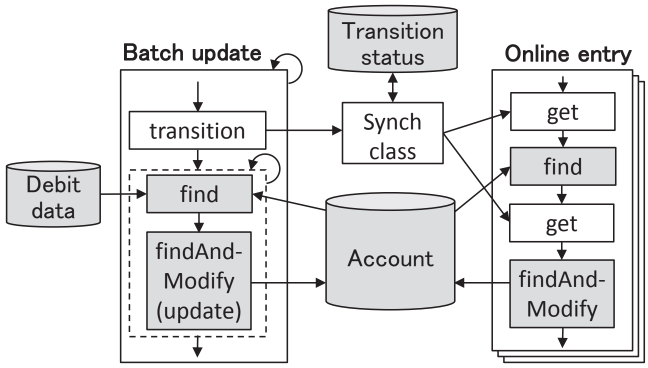



Figure 7: Software structure of prototype

completed, balance field is not changed and this document become the state shown in (4).

In this way, the processing stage transitions to "done", and we get the result (3) in the case of successful completion; we get (4) in the case of abnormal termination.

# 4 IMPLEMENTATION AND EVALUATIONS

## 4.1 Implementation of Prototype

To evaluate the proposed method, we constructed a prototype intending to manipulate the deposit accounts of the banking system. We use MongoDB Ver. 2.6.7 for the database; Java Ver. 1.6 for the programming language; MongoDB Java Driver Ver. 2.13 to access MongoDB from Java [11]. In addition, OS is Windows 7 (64bit). Figure 7 shows its software construction. The batch update and online entry programs are implemented by Thread class of Java to execute the both concurrently. Each program executes the following processes as shown in this figure: it query the data of the deposit account from the database (find); then, it updates the data of the database (findAndModify, update).

The batch update program executes the processing to debit from the deposit account collection (Account) in a lump-sum, based on the account and amount information stored in the debit data collection (Debit data). As shown in Fig. 4, the processing stage transitions from "initial" to "pending". This process is executed at the first (transition) of the batch update program (Batch update), then the batch update is executed. After its completion, the processing stage transitions to "applied", and the data in temp field is reflected into balance field. Incidentally, in the case of abnormal termination, it transitions to "rollback". Finally the processing stage transitions to "done". The information of the processing stage is stored in the transition status collection (Transition status), and it is accessed through "Synch class" by the batch update and online entry programs.

Meanwhile, the online entry executes the processing to deposit to each deposit account individually. As for this prototype, it was configured to perform deposits of certain amount of money from the plural terminals concurrently. Here, the online entry has to be accompanied by the OB update during the processing stage of the batch update is "applied". So, its program was configured to query the processing stage by Synch class (get). And, to query this data efficiently by the program without accessing the database, it is saved in the instance of Sync class, However, in the case where the processing stage transitions from one stage to the next stage during the online entry executing, there is the possibility of the incorrect OB update execution. In other words, in the case where the transition occurs between the "find" and "findAndModify" in Fig. 7, there may be the unnecessary OB update execution or the lack of it.

For this issue, We implemented Synch class using Synchronized keyword of Java, by which only one program can call it at the same time by the synchronization control. Then, we configured the online entry program to query the processing stage before not only "find" but also "findAndModify" as shown by the "get" in Fig. 7. And, in the case where the transition occurs between them, the online entry program performs a retry. Furthermore, in order to prevent the transition between "get" prior to findAndModify and the completion of findAndModify, we configured Synch class to wait a certain time before transition, which is requested by the batch update program. That is, the executing update of the online entry program can be completed before the transition by this way. Incidentally, while the processing stage is "applied", not only "balance" field but also the valid field has to be queried from this document. We implemented a class to manipulate the fields, and these manipulations were implemented by using the method of this class.

Since the online entries are executed from plural terminals concurrently, it is necessary to execute the concurrency control. So, we implemented the optimistic concurrency control by "last" field (update number) using findAndModify command, which is a method to perform the query and update of a document at the same time exclusively as mentioned in Section 2.2. And, in the case where the query condition matches to no data, the update is not performed and $null$ is returned as the query result. Therefore, we set the query condition of findAndModify command {"account":account number, "last": read updated number by "find" }, that is, the value of "last" is the result queried by find command just before. As a result, in the case where the target document was updated by the other program after the execution of this "find" command, no data matches this condition. And, in this case, the online entry program has to retry these processes from the beginning.

Table 1 shows the target fields at each processing stage, which is queried and updated. As for the batch update, it is not executed when the processing stage is "initial" or "done"; it updates the different fields from the online entry when the processing stage is "pending" or "rollback". So, there is no conflict between the batch update and online entry. However, when the processing stage is "applied", the both update the same fields: "balance", "last" and "temp". That is, there is the conflict between them. Therefore, as for the batch update program, we also implemented the optimistic concurrency control using findAndModify command similarly to the online entry program. Incidentally, the batch update program queries "balance" when the processing stage is "pending", which is updated by the online entry program at the same time. That is, there is conflict between them. However, as we already mentioned in Section 3, the case where the online entry is executed, the OB update result becomes valid, which is created based on the execution result of the former. In other words, the batch update result is not used. Therefore, the concurrency control for this query is not required.

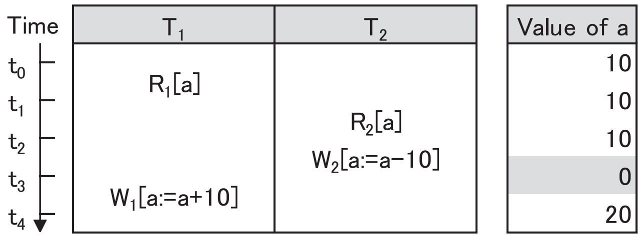



Figure 8: Lost update example of transaction

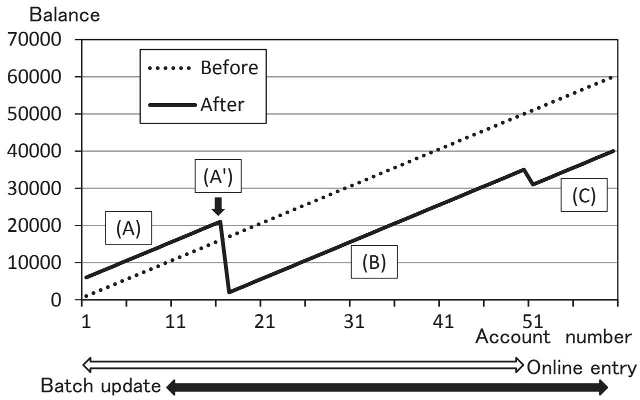



Figure 9: Result of case of successful completion

## 4.2 Evaluations of Concurrency Control

The proposal method does not lock the target documents through the duration of the batch update. That is, similar to the relational database, it has to be confirmed the inconsistencies by the concurrent execution of transactions do not occur. So, we performed the following three kinds of experiments to evaluate the concurrency control between the batch updates and online entry, using the prototype shown in Fig. 7.

First, we performed the experiment in the case of successful completion of the batch update. The purpose of this experiment is to confirm that there is no lost update occurred by the illegal interface between the batch updates and online entries. Figure 8 shows the example of the lost update, in which the time series manipulations on the data $a$ are executed by the transaction $T_1$ and $T_2$: $R_i$ indicates the query; $W_i$ indicates the update. And, the column "Value of a" shows the value of $a$ in the time series. As for the value of $a$, $T_1$ queries it by $R_1$ and updates by $W_1$; meanwhile, $T_2$ updates it by $W_2$. So, since the update result of $T_2$ is overwritten by $T_1$, it is lost. That is, the lost update has occurred.

In this experiment, as shown in Fig. 9, the number of the target deposit account is 60. And, its balance data is set prior to the experiment, which is calculated by the following equation as shown by the broken line.

$$balance = account\ number \times 1000 \tag{1}$$

Table 1: Read and write fields in each processing stage

| Processing stage | Batch update find | Batch update findAndModify | Online entry find | Online entry findAndModify |
|---|---|---|---|---|
| Initial | — | — | balance, last | balance, last |
| Pending | balance | temp.b | balance, last | balance, last, temp.o, temp.ob |
| Applied | last, temp | balance, last, temp (delete) | balance, last, temp | balance, last, temp (delete) |
| Rollback | — | temp (delete) | balance, last | balance, last |
| Done | — | — | balance, last | balance, last |

```
{ "account" : 15, "balance" : 18000, "last" : 3 }
{ "account" : 16, "balance" : 19000, "last" : 3 }
{ "account" : 17, "balance" : 20000, "last" : 3,
      "temp" : { "o" : 20000, "ob" : 0 } }
{ "account" : 18, "balance" : 21000, "last" : 3,
      "temp" : { "o" : 21000, "ob" : 1000 } }
{  "account" : 19, "balance" : 22000, "last" : 3,
      "temp" : { "b" : 0, "o" : 22000, "ob" : 2000 } }
{ "account" : 20, "balance" : 23000, "last" : 3,
      "temp" : { "b" : 1000, "o" : 23000, "ob" : 3000 } }
```

Figure 10: Data at end of pending

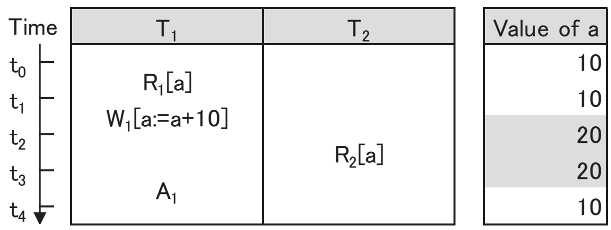



Figure 11: Dirty read example of transaction

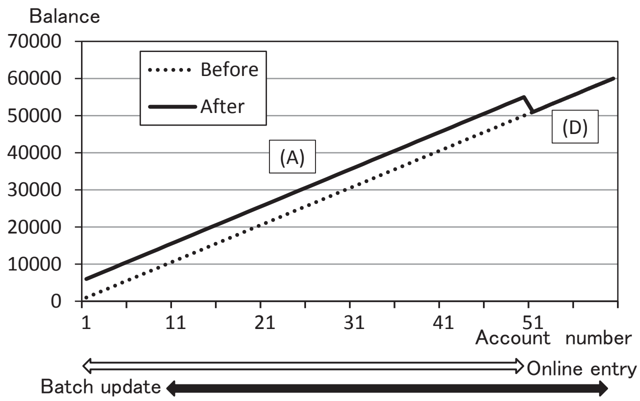



Figure 12: Result of rollback of batch update

Here, the horizontal axis shows the account number of the deposit account; the vertical axis shows its balance. Then, the batch update program debits 20000 from the deposit accounts which account number is between 11 and 60. Here, the account, which balance is less than 20000 at this debiting time, is excluded from this processing. In this experiment, since the batch update is successfully completed, the processing stage shown in Table 1 transitions from "pending" to "applied".

Meanwhile, the online entries are executed from five terminals concurrently, and each entry deposits 1000 to the deposit account which account number is between 1 to 50. Here, in order to avoid the conflict among the online entries, the different first update account number is assigned to each terminal: 1, 11, 21, 31, 41. Then, Each terminal updates the deposit account one after another. Here, after the program has processed the account which account number is 50, it processes the account which account number is 1. In this way, 50 deposit accounts are updated from each terminal.

The solid line in Fig. 9 shows this experimental result. The account indicated by (A) is not the target of the batch update or its balance was less than 20000. So, the batch update did not debit from it, and only the online entries deposited 5000. The account indicated by (B) is the target of the batch update, and the batch update or OB update debited 20000. Also, the online entries deposited 5000. So, the balance became 15000 reduction. The account indicated by (C) is not the target of the online entry, and only the batch update debited 20000. That is, even in the case where the batch update was executed concurrently with the online entry, no lost update occurs in both of the processes. As a result, we got the consistent update result.

Second, to investigate the change point from (A) to (B) in Fig. 9, which is shown by (A'), we performed the experiment, in which both of the online entry and batch update were interrupted when the processing stage transitioned to "applied". Incidentally, the other experimental environment is the same as the first experiment. Figure 10 shows the query results of the deposit account data at the end of the experiment, which is in the vicinity of (A').

Since the intermediate results of "temp" field at this time remained, the following data was queried. As for the account number 15 and 16, since the balance was less than 20000, neither of the batch update and OB update were performed. As for 17 and 18, since the balance was less than 20000 when the batch update was performed, the batch update was not performed. However, since the online entries deposit after this time, the balance exceeded 20000 and the OB update was performed. Lastly, as for 19 and 20, both of the batch update and OB update were performed. Incidentally, in the first experiment result, since the online entries were continued even after the transition to "applied", every balance of deposit accounts are grater than 0.

Third, we experimented the case of the abnormal termination of the batch update, and the processing stage was transitioned from "pending" to "rollback", which is shown in Table 1. The purpose of this experiment is to confirm that the dirty read of the online entry does not occur, even in the case of the abort and rollback of the batch update. Figure 11 shows the example of dirty read, and the representations are same as Fig. 8, and $A_1$ shows the abort of $T_1$. Though $T_1$ was aborted after updating $a$, $T_2$ had already queried this updated data. That is, since $T_2$ was executed using the data, which did not actually exist, the consistency of its result was not maintained.

We show the result of the third experiment in Fig. 12. Similar to Fig. 9, the broken line shows the balance data at the beginning of this experiment; the solid line shows the data at the end of this experiment. The former is the same as in the first experiment. In this experiment, the processing stage transitioned from "pending " to "rollback" to execute the rollback of the batch update, then "temp" field was deleted. As a result, as for the balance data in the range of (A), only the deposit of 5000 was executed by the online entries. On the contrary, the balance data does not change in the range of (D), which is outside of the online entry. Therefore, there was no dirty read of the online entries, and the consistency of the result was maintained. Therefore, the batch update could be canceled without affecting the online entry.

## 5 CONSIDERATIONS

First, we consider whether the ACID properties of the transaction are maintained by the proposal method. As for the atomicity, the batch update completes as either of the following state: its update results are queried after the processing stage transits to "applied" as shown in Fig. 9; it is canceled in the stage of "rollback" without affecting the online entry as shown in Fig. 12. Therefore, the consistency was also maintained, that is, the collection transitions from a consistent state to another consistent state.

Next, as for the isolation, the batch update updates the different fields from the online entry, and the intermediate results of each processing are not queried by the other when the processing stage is "pending" as shown in Fig. 3. Furthermore, in the both case of the successful completion and rollback, the batch update could be executed without affecting the online entries. So, the isolation is maintained. Lastly, as for the durability, the integration processing of the online entry and batch update results is executed when the processing stage is "applied". However, since the update results have been already reflected into the database, the durability is maintained by database management system of MongoDB. And, the query results do not change even if this process is interrupted.

Thus, the ACID properties of the transaction can be maintained by this method in MongoDB, even if the batch update is executed concurrently with the online entries. As shown in Section 3, this batch update corresponds to updating plural documents in a lump-sum. In other words, the update of plural documents in MongoDB can be executed as a transaction concurrently with the update of individual document.

Second, we consider the efficiency, which is the latency on the online entries by the batch update. As shown in Table 1, the batch update and online entry update the different fields from each other while the processing state is not "applied". So, the latency does not occur by the concurrency control for the conflict. On the other hand, since the both update the same fields, "balance" and "last", while the processing stage is "applied", the concurrency control is needed for these conflicts. Here, the concurrency control is executed by the optimistic concurrency control for not only the online entries but also each individual update of the batch update. Therefore, it can be executed without a long latency, such as waiting for the completion of the entire batch update.

In addition, when the processing stage transitions to "applied", the batch update itself was already completed and the online entry is executed using a valid data reflecting the batch and OB update results. So, for example, in the case where the batch update is not executed often, its processing in "applied" can be wait to execute until the frequency of the online entry becomes less comparatively. Incidentally, in this experiment, the delays of certain period of time were put in Sync class in order to prevent the transition of the processing stage during the updating of the online entry, by a simple way. And, this causes the latency of the other online entries. As for this issue, we consider it can be shortened by the immediate transition after the completion of this online entry.

Third, through our experiments, we found that the different concurrency control method from the relational database can be applied to MongoDB, which is a kind of document oriented NoSQL database. As for the relational database, the row lock method is generally used, and a key-range lock method supports the concurrency control for the wide range of rows of tables [3]. On the other hand, it is controlled as an update of the entire row even when a part of the row is updated. As for MongoDB, in contrast, although there is a restriction of the concurrency control to update plural documents collectively, there is no conflict among the update of the different fields. This shows, while the long-lived transaction is updating the large-capacity field such as videos, the other fields can be updated by the other transactions concurrently. In other words, though there is little necessity for the concurrency control in the NoSQL database now, we consider the different update model from the relational database will be necessary according with the spread of its application fields.

Lastly, the update of plural documents in a lump-sum is often executed in the actual business system operations. In particular, the case of the experiment in this paper is taken as a typical example [3]. On the other hand, it had been the problem in the MongoDB and other NoSQL databases. With the expansion of the application fields of the NoSQL databases, it is considered that the request for the data manipulation like this shall increase, which the lump-sum update is executed as the single transaction as well as the relational databases. So, we consider that this method is valid for such a data manipulation.

## 6 CONCLUSION

Recent years, the utilizations of the NoSQL databases are spreading. However, there is the problem that the plural objects cannot be updated with maintaining the ACID properties of transactions. In this paper, we proposed the lump-sum update method for MongoDB, which is a kind of NoSQL database. This method is based on the concept of the temporal update method to execute the batch update as the single transaction in the relational databases. Concretely, in this method, the results of the following update are stored in temporal fields of the document and only the valid data is queried: the batch updates; the online entry; the OB update, which is applied the batch update individually to the online entry result.

And, we showed that the plural documents can be updated as the single transaction in MongoDB by this method, even while the documents are being updated concurrently by the other transactions, that is, the online entries. Furthermore, we confirmed that this method achieves the above-mentioned function through the experiments using a prototype, which intended the deposit account.

Meanwhile, in the actual business systems, large number of transactions which update the plural data are executed concurrently. So, the future study will be focused on the concurrency control of such a update in the NoSQL databases.

## ACKNOWLEDGEMNTS

This work was supported by JSPS KAKENHI Grant Number 15K00161. Also, the motivation of this study is an extension of our method to the NoSQL database, which has aimed to execute the batch update as a transaction in the relational databases and has been registered as a patent [4]. We would like to appreciate the members of Mitsubishi Electric Information Systems Corp. who supported to get this patent.

## REFERENCES

[1] K. Banker, MongoDB in Action, Manning Pubns Co. (2011).
[2] C.J. Date, An Introduction to Database Systems, Addison-Wesley (2003).
[3] J. Gray, A. Reuter, Transaction Processing: Concept and Techniques, San Francisco: Morgan Kaufmann (1992).
[4] Mitsubishi Electric Information Systems Corp., T. Kudo, Database System, Japan patent 4396988 (2009).
[5] T. Kudo, Y. Takeda, M. Ishino, K. Saotome, and N. Kataoka, "Evaluation of Lump-sum Update Methods for Nonstop Service System," Int. J. of Informatics Society, Vol. 5, No. 1, pp. 21–28 (2013).
[6] T. Kudo, Y. Takeda, M. Ishino, K. Saotome, and N. Kataoka, "An implementation of concurrency control between batch update and online entries," Procedia Computer Science, Vol. 35, pp. 1625–1634 (2014).
[7] T. Kudo, Y. Takeda, M. Ishino, K. Saotome, and N. Kataoka, "Application of a Lump-sum Update Method to Distributed Database," Int. J. of Informatics Society, Vol. 6, No. 1, pp. 11–18 (2014).
[8] D. Laney, "3D Data Management: Controlling Data Volume, Velocity and Variety," META Group (2012) http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.
[9] H. Garcia-Molina, and K. Salem, "SAGAS," Proc. the 1987 ACM SIGMOD Int. Conf. on Management of data, pp. 249–259 (1987).
[10] MongoDB Inc., The MongoDB 3.0 Manual, http://docs.mongodb.org/manual/.
[11] MongoDB Inc., MongoDB API Documentation for Java, http://api.mongodb.org/java/.
[12] E. Redmond, and J.R. Wilson, Seven Databases in Seven Weeks: A guide to Modern Databases and The NoSQL Movement, Pragmatic Bookshelf (2012).
[13] P.J. Sadalage, and M. Fowler, NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence, Addison-Wesley Professional (2012).
[14] R. Snodgrass and I. Ahn, "Temporal Databases," IEEE COMPUTER, Vol. 19, No. 9, pp. 35–42 (1986).
[15] S.S. Sriparasa, JavaScript and JESON Essentials, Packt Pub. Ltd. (2013).
[16] B. Stantic, J. Thornton and A. Sattar, "A Novel Approach to Model NOW in Temporal Databases," Proc. 10th Int. Symposium on Temporal Representation and Reasoning and Fourth Int. Conf. on Temporal Logic, pp. 174–180 (2003).

(Received September 29, 2015)
(Revised December 8, 2015)

**Tsukasa Kudo** received the M.E. from Hokkaido University in 1980 and the Dr.Eng. in industrial science and engineering from Shizuoka University, Japan in 2008. In 1980, he joined Mitsubishi Electric Corp. He was a researcher of parallel computer architecture, an engineer of application packaged software and business information systems. Since 2010, he is a professor of Shizuoka Institute of Science and Technology. Now, his research interests include database application and software engineering. He is a member of IEIEC and Information Processing Society of Japan.

**Masahiko Ishino** received the master's degree in science and technology from Keio University in 1979 and received the Ph.D. degree in industrial science and engineering from graduate school of Science and technology of Shizuoka University, Japan in 2007. In 1979, he joined Mitsubishi Electric Corp. From 2009 to 2014, he was a professor of Fukui University of Technology. Since 2014, he belong to Bunkyo University. Now, His research interests include Management Information Systems, Ubiquitous Systems, Application Systems of Data-mining, and Information Security Systems. He is a member of Information Processing Society of Japan, Japan Industrial Management Association and Japan Society for Management Information.

**Kenji Saotome** received the B.E. from Osaka University, Japan in 1979, and the Dr.Eng. in Information Engineering from Shizuoka University, Japan in 2008. From 1979 to 2007, he was with Mitsubishi Electric Corp., Japan. Since 2004, he has been a professor of Hosei business school of innovation management. His current research areas include LDAP directory applications and single sign-on system. He is a member of the Information Processing Society of Japan.

**Nobuhiro Kataoka** received the master's degree in electronics from Osaka University, Japan in 1968 and the Ph.D. in information science from Tohoku University, Japan in 2000. From 1968 to 2000, he was with Mitsubishi Electric Corp. From 2000 to 2008, he was a professor of Tokai University in Japan. He is currently the president of Interprise Laboratory. His research interests include business model and modeling of information systems. He is a fellow of IEIEC.

# Dissolve in Scents Using Pulse Ejection When Combinations of Scents Were Changed

Sayaka Matsumoto*, Shutaro Homma*, Eri Matsuura*, Shohei Horiguchi*, and Ken-ichi Okada**

* Graduate School of Science of Technology, Keio University, Japan
** Faculty of Science and Technology, Keio University, Japan
{matsumoto, honma, matsuura, shohei, okada}@mos.ics.keio.ac.jp

***Abstract*** - A trial to raise a sense of reality by using scents with various types of media has lately attracted much attention. In addition, it is thought that we can raise a sense of reality more by not only adding scents, but also expressing the movement of scents with that of the picture. We aimed at the development of the presentation technique to express dissolve in scents with paying attention to changing the intensity of two types of scents. The results of experiments revealed that receivers may feel dissolve by presenting fade-in and fade-out in scents which are overlapped in three breathes. Then we examined whether they can feel dissolve in scents in different types of scents. It is expected that the technique can raise realistic sensations when scents are presented in accordance with pictures by establishing the technique of dissolve in scents.

***Keywords***: Olfactory display, Pulse ejection, Dissolve in scents

## 1 INTRODUCTION

Information transmission and communication tends to be limited to visual information and audio information. However, the transmission of information via all five senses (sight, hearing, touch, smell and taste) has lately attracted much attention [1]. Olfactory information recognized by the olfactory organs differs from the information recognized via the other four senses [2]. The sense of smell powerfully affects humans since olfactory information is directly transmitted to the cerebral limbic system that governs emotions. In addition, olfactory information has high importance since it is thought that the presentation of olfactory information is effective as a means to enhance the sense of reality like three-dimensional vision and sound [3].

For transmitting scents together with other media, it is necessary to control the presentation of scent in accordance with the changes in images/sounds over time. In doing so, it is more effective to enhance the sense of reality. Therefore, we paid attention to both the change of types and the intensity of scents and developed presentation techniques in scents changing the intensity of two types of scents. Among these, this study presents a technique in scents that enables the receivers to feel “dissolve”, which we defined as “the second scent becomes gradually strong at the same time as the first scent becomes gradually weak”.

First, we constructed presentation techniques in scents that the receivers to feel “fade-out in scents” (the scent becomes gradually weak) and “fade-in in scents” (the scent becomes gradually strong) using pulse ejection. Then, we expressed dissolve in scents with combining fade-out in scents and fade-in in scents, and examined whether the receivers can feel dissolve in scents.

## 2 RELATED WORK

### 2.1 The Study of Adding Scents

Trials on the transmission of olfactory information together with other media are currently being conducted. “Scents of Space” which Haque et al. developed is the art work of lights and scents, which can carry a scent with the wind from one wall of the room and send it to receivers [4]. A trial to present scents in accordance with movie at a movie theater carried out using a device called Aromageur [5] which is the scent generator can save the recipe of the scents [6]. These trials are aimed for the enhancing the sense of reality by adding scents in the room or with videos.

There are many scenes in videos and TV programs, such as the scene which many smelling objects appear at the same time, suddenly appear and gradually disappear. Therefore, it is necessary not only to present scents but also to control the presentation of scent in accordance with the changes in images/sounds over time. In doing so, it is more effective to enhance the sense of reality.

However, they paid attention only to adding scents with other media in related works and the study that are paid attention to changes of types and the intensity of scents are not performed so many. The conventional presentation method of scents which are used in related works continues emitting scent at high density for a long time that the receivers can feel enough scents. Too much scent emitted over a continuous period leaves in the air. If the scent presented before mixes with scent presented later, there is a possibility that the receivers cannot feel the change of the types of scents.

Moreover, too much scent causes human adaptation to the scent, and thus, the receivers may not feel the intensity of scents properly. From such problems, it was possible to add scents but it was difficult to change types and the intensity of scents.

### 2.2 Presentation Technique That Emit Scent for Very Short Periods of Time

In our previous research, we minimized the influence of the scents to spread in the air by emitting scents for just very short periods of time and reduced the fragrance which remained in the air. In this way, we can reduce human adaptation and change scents without scents being mixed. We defined this presentation technique that emits scent for very short periods of time as “pulse ejection” [7], and we studied about the change of kind of scents and the intensity of the scents by using pulse ejection. About the change of the kind of scents, for example, we measured the interval of the ejection time that human can recognize two types of scents clearly without being mixed, and developed a presentation technique that the receivers can feel two types of scents in one breath [8]. Furthermore, applying this presentation technique, we developed presentation techniques that the receivers can feel the strength of the relation between two types of scents by presenting a weak scent earlier and a strong scent later [9]. Besides, when we paid attention to the change of kind of scents every fixed time, it was revealed that the receivers can feel the change of kind of scents every two breathes [10]. About the change of the intensity of scents, for example, we developed presentation techniques that the receivers can feel the scent is coming near or going away by changing the intensity of a scent every two breathes [11]. Like these, we enabled to present the change of scents by using pulse ejection.

# 3 DISSOLVE IN SCENTS

## 3.1 Presentation Technique of Dissolve in Scents

As it was previously mentioned at Section 2.2, we studied about the change of types and the intensity of scents independently to express the presentation of scents by using pulse ejection. In this study, we paid attention to both the change of types and the intensity of scents and devised presentation techniques in scents changing the intensity of two types of scents.

For transmitting scents together with video, it is thought that it is effective to use the presentation technique of scents that is suitable for scene conversions of the videos. Therefore, we focus on dissolve in scene conversions of the videos. In videos, dissolve is a scene conversion that “the next scene is gradually superimposed as the former scene fades out” [12]. We propose to express dissolve in scents based on dissolve in videos. This study presents a technique in scents that enables the receivers to feel “dissolve”, which we defined as “the second scent becomes gradually strong at the same time as the first scent becomes gradually weak”. To develop the presentation technique to express dissolve in scents, it is necessary to develop the presentation technique that the receivers can feel the scent becomes gradually strong and gradually weak. Therefore, we defined “fade-in in scents” as the presentation technique that the scent becomes gradually strong, and “fade-out in scents” as the presentation technique that the scent becomes gradually weak. Then, we present dissolve in scents with combining fade-out in scents and fade-in in scents.

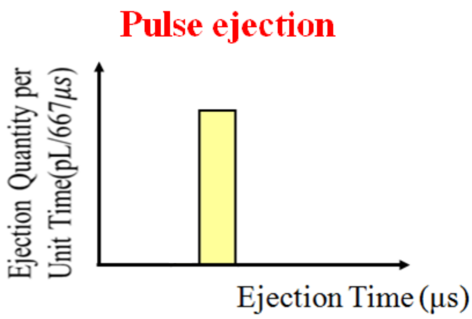



Figure 1: Image of pulse ejection

## 3.2 Pulse Ejection

When we present dissolve in scents, if we use the conventional presentation method of scents that emit scent at high density for a long time that the receivers can feel enough scents, the scent presented before mixes with scent presented later and it is difficult to feel scents as we expected. Besides, the scent which was presented for a long time causes human adaptation to the scent. The receivers may not feel the intensity of scents properly. In this study, we propose the scent presentation technique to create dissolve by using pulse ejection. An olfactory display we developed uses the technique used in ink-jet printer, and can use pulse ejection.

Pulse ejection is controlled scents as quantity by two parameters, ejection quantity of scent per unit time and ejection time (Fig 1). Besides, this device can change the ejection time at 667 μs intervals, so that it can present scent during only one breath. In presenting infinitesimal quantity of scent like this, pulse ejection can minimize the lingering of scents in the air.

# 4 EXPRESSION OF FADE-IN AND FADE-OUT IN SCENT

In this chapter, to determine how many times we should change the intensity of scents to express fade-in and fade-out at first. We examine how many times human can feel the change of the intensity of scents(Preliminary Experiment). After that, we developed the presentation techniques to express fade-in in scents and fade-out in scents.

## 4.1 Olfactory Display

We developed an olfactory display called “Fragrance Jet 2(FJ2)”. Figure 2 shows the condition of this experiment by FJ2. In this study, we used FJ2 which is a model to put in front of a participant. However, when we actually present scents together with video, it is more effective to use a mobile display that a participant puts on, so that it won't prevent viewing video.

This display uses the technique used in ink-jet printer in order to produce a jet which is broken into droplets from the small hole in the ink tank. This device can use pulse ejection for scent presentation so that the issue such as scent

Figure 2: The condition of this experiment by Fragrance Jet 2(FJ2)

lingering and care to eject scent can be minimized. The display can set up one scent ejection head. This head can store three small tanks and one large tank, thus this display can contain 4 types of scents. There are 127 minute holes in the head connected to the small tank and 256 minute holes in the head connected to the large tank. Moreover, the display can emit scent from multiple holes at the same time. We denote the number of minute holes emitting at one time as "the number of simultaneous ejections". So, the ejection quantity is adaptable to 0-127 (small tank), 0-256 (large tank) if the ejection time is set. In this study, the pulse ejection time is set to 100msec, so the ejection quantity is controlled by the change of the number of simultaneous ejections. We define the number of simultaneous ejections as the "intensity" of scents at this display.

## 4.2 Preliminary Experiment : The Number of Times of the Change

### 4.2.1 Experimental Method

At first, we measured the detection threshold of each participant. The detection threshold is the smallest density at which scent can be detected and where the user does not need to recognize the kind of a smell. The experiment to determine the detection threshold was conducted using the scent of banana stored in a large tank. The intensity was changed by 10 between 10 and 250. We use the triangle test [13] to judge the detection threshold in the measurement. In the triangle test, three stimuli are presented at random, where one of them is scented and the other two are odorless. The participant then answers when the scented odor was presented. Furthermore, we used the raising method (the first intensity was 10) to measured the detection threshold. The detection threshold was determined by the intensity which the participant answered correctly twice in a row. If the participant selected the wrong answer, the intensity was raised by 10.

After measuring the detection threshold, we examined how many times the participant could feel a change of the intensity of scent when the intensity was changed by 10 between the detection threshold of each participant and 250. We presented two scented ejections of different intensity to each participant. The interval between two scented ejections was 4 seconds. We presented first scented ejection in first breath and presented second scented ejection in second breath. We signaled the timing of breathing by sounds. When the scent is ejected, the countdown starts with the auditory cue. Scent emission then commences 0.5 sec after giving the cue "Go" according to previous study[14].

When we examine the number of times that the participant can feel the change of intensity of scent, we used two methods, rising method and dropping method. In rising method, we started the experiment from the intensity of detection threshold to 250. We prepared a reference value and a comparison value. The first reference value is the intensity of detection threshold and we started the experiment from the reference value. We presented two scented ejections to each participant in random order. One scented ejection was presented in the intensity of reference value and the other one was presented in the intensity of the comparison value which is larger than the reference value by 10. Then, we instructed the participant to answer which of the two was strong. If the participant answered correctly twice in a row, we judged that the participant can distinguish the intensity of the two, and recorded the comparison value. After that, we substituted the intensity of the comparison value for the next reference value and resumed the experiment. If the participant selected the wrong answer, we changed the comparison value to the value that is larger than the last comparison value by 10 and resumed the experiment. When the comparison value was reached 250(the maximum value), we finished the experiment. In dropping method, we started the experiment from 250 to the intensity of detection threshold. The same as the rising method, we prepared a reference value and a comparison value. The first reference value is 250 and we started the experiment from the reference value. We presented two scented ejections to each participant in random order. One scented ejection was presented in the intensity of reference value and the other one was presented in the intensity of the comparison value which is smaller than the reference value by 10. Then, we instructed the participant to answer which of the two was weak. If the participant answered correctly twice in a row, we judged that the participant can distinguish the intensity of the two and recorded the comparison value. After that, we substituted the intensity of the comparison value for the next reference value and resumed the experiment. If the participant selected the wrong answer, we changed the comparison value to the value that is smaller than the last comparison value by 10 and resumed the experiment. When the comparison value was reached the detection threshold of each participant (the minimum value), we finished the experiment.

In rising method, participants were 5 men and 3 women in their 20s. In descending method, participants were 4 men and 3 women in their 20s. Participants of the descending method were same as those of the rising method.

### 4.2.2 Results

The detection threshold of banana was 21±11.7 (average ± standard deviation). Figures 3 and 4 show the results of the

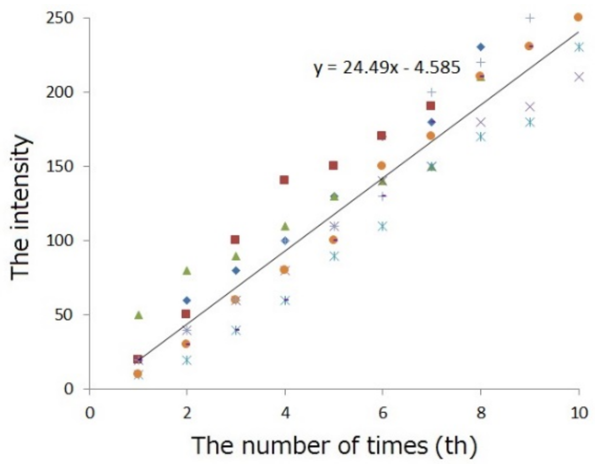



Figure 3: The intensity of scents that participants could notice (rising method)

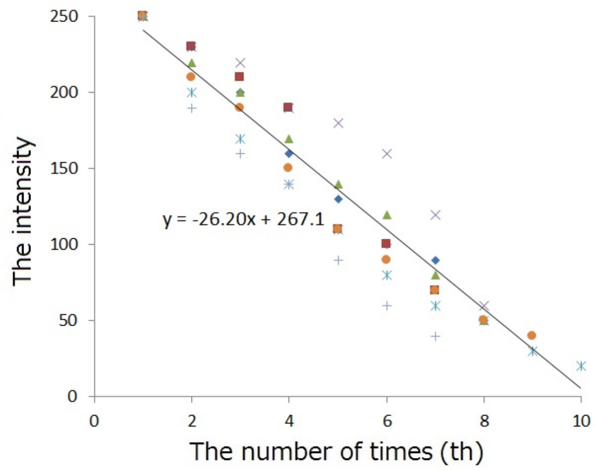



Figure 4: The intensity of scents that participants could notice (dropping method)

Table 1: Calculated values (rising method)

| The number of times | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Intensity | 20 | 45 | 70 | 95 | 120 | 145 | 170 | 195 | 220 | 245 |

Table 2: Calculated values (dropping method)

| The number of times | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Intensity | 247 | 220 | 193 | 166 | 139 | 112 | 85 | 58 | 31 | 4 |

number of times that participants felt the changes of the intensity in case of the rising method and the dropping method. The horizontal axis shows the number of times that participants could feel the changes of the intensity. The vertical axis shows the intensity of the scent that participants could notice. We plotted the intensity for each participant, such as the first change of the intensity that he or she was able to feel, and the second change of the intensity that he or she was able to feel. It is revealed that all participants could notice the changes of the intensity more than 7 times both in rising method and in dropping method. Furthermore, it is thought that the results of Fig. 3 and 4 can be approximated by a straight line. Fechner's law says subjective sensation increases proportional to the logarithm of the stimulus intensity. However, in this experiment, we checked whether participants could feel changes of the intensity of slightly small amount of scent by using pulse ejection. So we think that the result might become linearly. We calculate the average value of the intensity for each number of times and showed the approximation straight line to these graphs. Tables 1 and 2 show the values of the intensity of scent that are calculated by substituting the number of times for the approximation and rounded off to the first decimal place. Using these values, we developed the presentation techniques to express fade-in in scents and fade-out in scents.

## 4.3 Fade-in and Fade-out in Scents

Using the values of Tables 1 and 2, we examined whether the participant can feel fade-in in scents and fade-out in scents when we changed the intensity every 1 breath. When we presented one of the values of Table per breath to each participant and tested whether he/she can feel fade-in and fade-out. Participants were 5 people in their 20s. As a result, in fade-in, all participants answered that they felt “the scent becomes gradually strong”. However, most of them were not able to feel scent when the intensity of the scent was less than 45, and were not able to feel the change of the intensity when the intensity of the scent was greater than 170. From this, participants could feel six changes of the intensity. In fade-out, all participants answered that they felt “the scent becomes gradually weak”. However, most of them were not able to feel scent when the intensity of the scent was less than or equal to 31, and were not able to feel the change of the intensity when the intensity of the scent was greater than or equal to 193. From this, participants could feel six changes of the intensity.

Therefore, we decided the number of the changes of the intensity up to six times. We decided to present the intensity of scents in the range of 45 – 170 in fade-in. We decided to present the intensity of scents in the range of 35 – 170 in fade-out to make the maximum value of fade-out same to that of fade-in.

### 4.3.1 Experimental Method

We examined how short fade-in and fade-out in scents can be felt. As described above, we defined fade-in in scents as the presentation technique that the scent becomes gradually strong, and fade-out in scents as the presentation technique that the scent becomes gradually weak. In addition, we judged that participant can feel the change of the intensity smoothly when they can feel a lot of changes of the intensity. The experiment was conducted using the scent of banana stored in a large tank. In fade-in, we divided the intensity

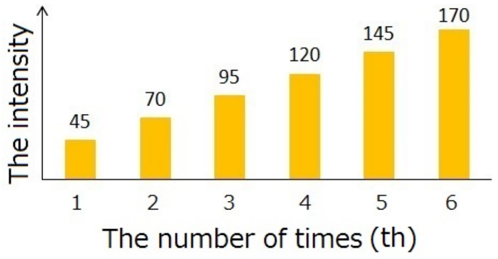



Figure 5: Image of (a) 6 phases (fade-in)

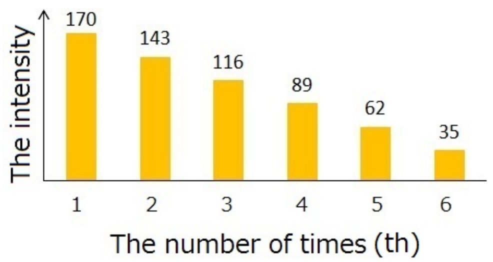



Figure 6: Image of (e) 6 phases (fade-out)

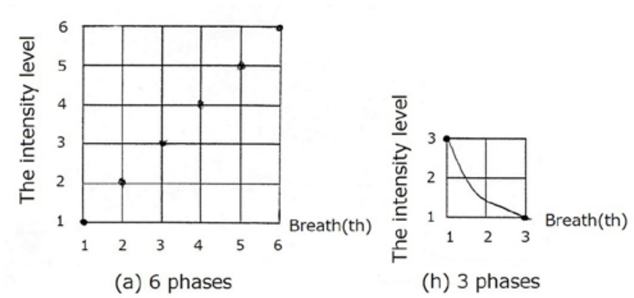



Figure 7: Example in a way of answering graphs

between minimum 45 and maximum 170 linearly into (a) 6 phases, (b) 5 phases, (c) 4 phases and (d) 3 phases. In fade-out, we divided the intensity between minimum 35 and maximum 170 linearly into (e) 6 phases, (f) 5 phases, (g) 4 phases, (h) 3 phases. As examples, Fig. 5 and 6 show the images of presentation of scent about (a) 6 phases and (e) 6 phases. The horizontal axis shows the number of times and the vertical axis shows the intensity of the scent that we presented. We presented each scented ejection once per 1 breath, and numbers in Fig. 5 and 6 show the intensity of scents that we presented for each breath. The interval between each breath was 4 seconds. The same as preliminary experiment, we signaled the timing of breathing by sounds. After we presented 1 series of phases, we asked the participant how they felt about the changes of the intensity of scent. For example, when we presented (a) 6 phases in fade-in, we presented 6 shots in a row. After taking a short break, we presented another 1 series of phases to the participant. We instructed them to answer by 2 ways, multiple choices and graphs, when they answered. In a way of answering by multiple choices, the participant was instructed to choose the most suitable item from seven evaluation items as follows; "The

Table 3: The number of participants that answered "The scent becomes gradually strong"

| Fade-in | (a) 6 phases | (b) 5 phases | (c) 4 phases | (d) 3 phases |
|---|---|---|---|---|
| The number of participants | 12 | 7 | 8 | 8 |

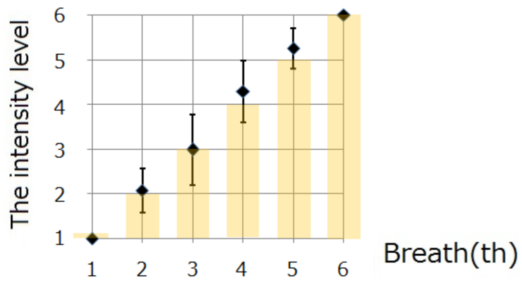



Figure 8: The result of the graphs in (a) 6 phases

scent becomes strong suddenly", "The scent becomes gradually strong", "The scent becomes weak suddenly", "The scent becomes gradually weak", "There was no change in feeling the scent", "There was no scents as a whole","The scent irregularly changed from feeling strong or weak". In a way of answering by graphs, we used plotting paper like Fig. 7. Figure 7 also shows the example of drawing and plotting. The horizontal axis shows the number of times and the vertical axis shows the intensity level of scent. Level 1 is the weakest, and Level maximum ((e) 6 phases: level 6, (h) 3 phases: level 3) is the strongest. In fade-in ((a) 6 phases - (d) 3 phases), we set the intensity in first breath to level 1. In fade-out ((e) 6 phases - (h) 3 phases), we set the intensity in first breath to level maximum. Before beginning this experiment, we presented the scent of level 1 and level 6 to each participant in advance and instructed them to memorize the sense of the intensity of these levels. We instructed each participant to draw or plot the changes of the intensity of scents that he or she felt after second breath based on these levels. Participants were 10 men and 6 women in their 20s.

### 4.3.2 Results and Considerations

At first, we describe the result and the consideration of fade-in in scents. Next, we describe the result and the consideration of fade-out in scents.

- **Fade-in in Scents**

First, we describe the results about a way of answering by multiple choices. We judged the most suitable phases to create fade-in in scents with the number of participants that answered "The scent becomes gradually strong". Table 3 shows the number of the answers for four types of phases to create the impression of "gradually strong". It is revealed that the largest number of participants answered "gradually strong" in (a) 6 phases. The result of fade-in in multiple choices indicates that the participants could feel fade-in the best in (a) 6 phases.

Next, we describe the results about a way of answering by graphs. We judged that participants could feel the changes of the intensity smoothly when the number of the changes of the intensity participants could feel was the largest. We plotted the average values that are calculated from the values that each participant drew in each breath in each phase. For example, Fig. 8 shows the plotting points of (a) 6 phases. The horizontal axis shows the number of times and the vertical axis shows the intensity level of scent. The bar graph expresses the intensity of the scent that we presented. To examine whether the participants could feel the change of the intensity of scent between contiguous two breaths, such as the first breath and the second breath, the average values between contiguous two breaths were analyzed using t-test. If there was no significant difference between contiguous two breaths, the average values between n-th breath and (n+2)-th breath were analyzed. Significant differences were found in all contiguous two breaths at (a) 6 phases, (c) 4 phases and (d) 3 phases ($p < 0.05$). The result of these indicates that the participants could feel all changes of intensity of scent that we presented in those three phases. That is, the participants could feel six changes of the intensity in (a) 6 phases, four changes of the intensity in (c) 4 phases, three changes of the intensity in (d) 3 phases. On the other hand, in (b) 5 phases, it follows that participants could feel four changes of the intensity. The result of fade-in in graphs indicates that the participants could feel fade-in the most smoothly in (a) 6 phases. In light of results of two ways above, the participants were likely to feel fade-in in scents the best in (a) 6 phases.

## ● Fade-out in Scents

First, we describe the results about a way of answering by multiple choices. We judged the most suitable phases to create fade-out in scents with the number of participants that answered “The scent becomes gradually weak”. Table 4 shows the number of the answers for 4 types of phases to create the impression of “gradually weak”. It is revealed that the largest number of participants answered “gradually weak” in (e) 6 phases. The result of fade-out in multiple choices indicates that the participants could feel fade-out the best in (e) 6 phases.

Next, we describe the results about a way of answering by graphs. We judged that participants could feel the changes of the intensity smoothly when the number of the changes of the intensity participants could feel was the largest. We plotted the average values that are calculated from the values that each participant drew in each breath in each phase. For example, Fig. 9 shows the plotting points of (e) 6 phases. The horizontal axis shows the number of breaths and the vertical axis shows the intensity level of scent. The bar graph expresses the intensity of the scent that we presented. To examine whether the participants could feel the change of the intensity of scent between contiguous two breaths, such as the first breath and the second breath, the average values between contiguous two breaths were analyzed using t-test. If there was no significant difference between follows that participants could feel three changes of the intensity. In (h) 3 phases, it follows that participants could feel two changes of

Table 4: The number of participants that answered “The scent becomes gradually weak”

| Fade-out | (e) 6 phases | (f) 5 phases | (g) 4 phases | (h) 3 phases |
|---|---|---|---|---|
| The number of participants | 12 | 6 | 6 | 5 |

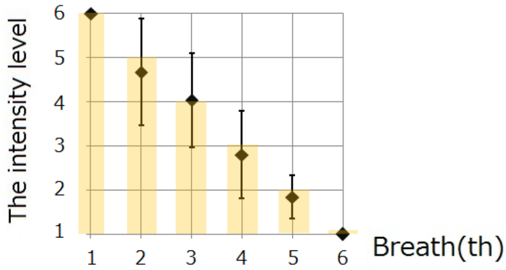



Figure 9: The result of the graphs in (e) 6 phases

the intensity. The result of fade-out in graphs indicates that the participants could feel fade-out the most smoothly in (e) 6 phases. In light of results of two ways above, the participants were likely to feel fade-out in scents the best in (e) 6 phases.

From these results, we combined (e) 6 phases in fade-out and (a) 6 phases in fade-in, and examined whether the participants can feel dissolve in scents.

# 5 EXPRESSION OF DISSOLVE IN SCENTS

We defined dissolve in scents as the presentation technique that “the second scent becomes gradually strong as the first scent becomes gradually weak” and “two types of scents were felt in a single breath”. In addition, we judged that participant can feel the change of the intensity smoothly when they can feel a lot of changes of the intensity. That is, there are the parts that two types of scents are presented in one breath to express dissolve in scents. In this experiment, we examined how many overlapping parts the presentation technique have that participants can feel dissolve the best when we presented the overlapping part that two types of scents are presented once per one breath.

## 5.1 Experimental Method

The experiment was conducted using the scent of banana stored in a large tank to express fade-out in scents and the scent of mint stored in a small tank to express fade-in in scents. In fade-out, we changed the intensity of scents six times using (e) 6 phases in the experiment 4.3. In fade-in, as the size of the tank was different from the experiment 4.3, we converted the intensity of scent (stored in a large tank) of (a) 6 phases in the Experiment 4.3 into those of scent stored in a small tank. Additionally, we measured the detection threshold of mint using the similar method to measure the detection threshold and similar participants in Section 4.2.1.

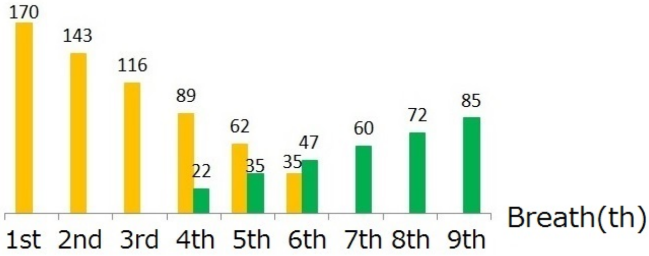



Figure10: Image of overlapping (“Dissolve 3”)

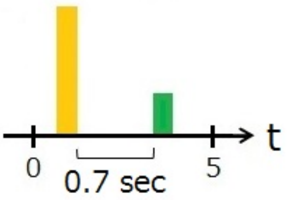



Figure 11: A presentation of the overlapping part

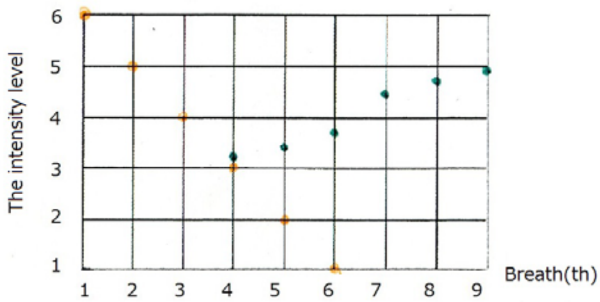



Figure 12: Example in a way of answering graphs (“Dissolve 3”)

As a result, the detection threshold of mint was 14±7.3 (average ± standard deviation). In this experiment, we prepared five patterns, from “Dissolve 1” to “Dissolve 5”. For example, “Dissolve 1” is a presentation that there is one overlapping part at which two types of scents are presented in one breath, and “Dissolve 5” is a presentation that there are five overlapping parts. Figure 10 shows the image of a presentation of “Dissolve 3”. The “Dissolve 3” has three overlapping parts, so there are nine breaths overall in “Dissolve 3”. In Fig. 10, yellow shows the scent of banana, and green shows the scent of mint. Numbers in Fig. 10 shows the intensity of scents that we presented for each breath. When combining fade-out and fade-in, we used a presentation technique that the receivers can feel two types of scents in one breath [8]. Figure 11 shows a presentation of the overlapping part. The horizontal axis shows the breaths and the vertical axis shows the intensity of the scent. The interval between two types of scents was set to 0.7sec [8]. We presented the scent of banana before and the scent of mint after, in spite of the intensity of the scents. The interval between each breath was 4 seconds in the Experiment 4, for smelling two types of scents, the interval between each breath was set to 5 seconds in this experiment. The same as

Table 5: The number of participants that answered “I felt two types of scents in “Dissolve n”

| Dissolve | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| The number of participants | 15 | 15 | 16 | 15 | 15 |

Table 6: The number of participants that answered “I felt two types of scents in one breath”

| Dissolve | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| The number of participants | 8 | 10 | 9 | 10 | 9 |

Table 7: The number of participants that could feel dissolve

| Dissolve | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| The number of participants | 8 | 7 | 8 | 7 | 7 |

preliminary experiment, we signaled the timing of breathing by sounds. After we presented one “Dissolve” to the participant, we asked the participant how they felt. After taking a short break, we presented another “Dissolve” to the participant. We instructed them to answer by two ways, sentences and graphs, when they answered. In a way of answering by sentences, the participant was instructed to answer about the “types” and the “intensity” of scents. In a way of answering by graphs, we used plotting paper like Fig. 12. The horizontal axis shows the number of times and the vertical axis shows the intensity level of scent. Level 1 is the weakest, and Level 6 is the strongest. The numbers of breath were different for each “Dissolve”, so we changed the size of plotting paper for each “Dissolve”. Figure 12 shows the example of “Dissolve 3”. We set the intensity in first breath to level 6. Before beginning this experiment, we presented the scent of level 1 and level 6 of each scent to each participant in advance and instructed them to memorize the sense of the intensity of these levels. We instructed each participant to draw or plot the changes of the intensity of scents that he or she felt after second breath. At the time, the participant was instructed to change the color of the pen when he/she felt some types of scents. Figure 12 also shows the example of plotting. Participants were 10 men and 6 women in their 20s.

## 5.2 Results and Consideration

First, we describe the results about a way of answering by sentences. In questions about the types of the scents, we asked participants “How many scents do you feel in “Dissolve n((n) is integers, $1 \leqq n \leqq 5$)” ?” Table 5 shows the number of the participants that answered “I could feel two types of scents in “Dissolve n”.” for five types of Dissolves. Table 5 indicates that about 80% of participants could feel two types of scents. There were few participants that felt only one kind of the scent because they could not notice the changes of scents, and felt three types of scents because the scents were mixed. Also, in questions about the types of the scents, we asked participants “Can you feel two types of scents in one breath?" Table 6 shows the number of the participants that answered "I could feel two types of scents in one breath." for five types of Dissolves. Table 6 indicates that more than half of participants could feel two types of scents without mixed in one breath from "Dissolve 1" to "Dissolve 5". We judged that the participants who answered "I felt two types of scents in one breath" (the question about the types of scents) and "I felt the first scent becomes gradually weak and the second scent becomes gradually strong" (the question about the intensity of scents) could felt dissolve. Table 7 shows the number of the answers for five types of "Dissolve" to create the impression of dissolve. It is revealed that the half number of participants could feel dissolve in scents in all "Dissolve"s. Even out of those, the number of participants that could feel dissolve in scents at "Dissolve 1" and "Dissolve 3" was the largest.

Next, we describe the results about a way of answering by graphs. We judged that participants could feel the changes of the intensity smoothly when the number of the changes of the intensity participants could feel was the largest. We plotted the average values that are calculated from the values that each participant drew in each breath in each "Dissolve". For example, Fig. 13 shows the plotting points of "Dissolve 3". The horizontal axis shows the number of breaths and the vertical axis shows the intensity level of scent. The level which participants could not feel scents was set to the level 0. The bar graphs express the intensity of the scent that we presented. In "Dissolve 3", the scent of banana was presented in between the first breath and the sixth breath and the scent of mint was presented in between the fourth breath and the ninth breath. That is, we presented two types of scents in between the fourth breath and the sixth breath. The bar graph expresses the intensity of the scent that we presented.

To examine whether the participants could feel the change of the intensity of scent between contiguous two breaths, such as the first breath and the second breath, the average values between contiguous two breaths for each scent were using t-test. If there was no significant difference between contiguous two breaths, the average values between n-th breath and (n+2)-th breath for each scent were analyzed.

In this paper, we describe the result about "Dissolve 3". As a result of comparison in the scent of banana at "Dissolve 3", significant differences were found in between the first breath and the second breath, the second breath and the third breath, the third breath and the fourth breath and the fourth breath and the sixth breath ($p < 0.05$). The result of this indicates that the participants could feel changes of intensity of scent at the first, second, third, fourth and sixth breath. That is, the participants could feel five changes of the intensity in the scent of banana at "Dissolve 3". As a result of comparison in the scent of mint at "Dissolve 3", significant differences were found in between the fourth breath and the sixth breath, the sixth breath and the seventh breath and the the intensity in the scent of mint at "Dissolve 3". The similar comparison was also conducted for other "Dissolve"s, and it was found that the number of changes of the intensity that participants could feel was the largest in "Dissolve 3". Therefore, the result of dissolve in graphs indicates that the participants could feel dissolve the most smoothly in "Dissolve 3". In light of results of two ways above, the partici

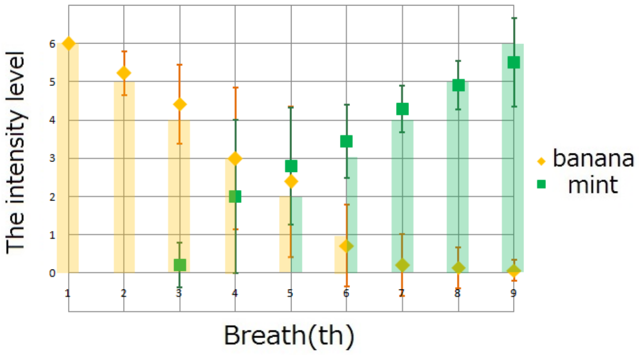



Figure 13: The result of the graphs in "Dissolve 3"

pants likely to feel dissolve in scents the best at "Dissolve 3" in those five presentation methods.

### 5.3 Expression of Dissolve in Differences in Kind of the Scents

Based on Experiment 5.1, we examined whether the participant can feel dissolve in scents in different types of scents in this sections.

#### 5.3.1 Combination of the Scents

We used the presentation technique of "Dissolve 3" and examined whether the participants could feel dissolve when we changed the types of scents. The experiment was conducted using the scent of banana stored in a large tank, and the scent of mint, rose and lavender stored in a small tank. In fade-out, we changed the intensity of scents six times using (e) 6 phases in the experiment 4.3 at a large tank. As the size of the tank was different from the experiment 4.3, we converted the intensity of scent (stored in a large tank) of (e) 6 phases in the Experiment 4.3 into those of scent stored in a small tank. In fade-in, we changed the intensity of scents six times using (a) 6 phases in the experiment 4.3 at a large tank. As the size of the tank was different from the experiment 4.3, we converted the intensity of scent (stored in a large tank) of (a) 6 phases in the Experiment 4.3 into those of scent stored in a small tank. Additionally, we measured the detection threshold of lavender and rose using the similar method to measure the detection threshold and similar participants in Section 4.2.1. As a result, the detection threshold of lavender was 11±3.2 and the detection threshold of rose was 16±5.2 (average ± standard deviation).

We prepared six patterns of combination of the scents, "banana / rose", "banana / lavender", "rose / banana", "rose / mint", "lavender / banana" and "lavender / mint". In terms of notation, for example, "banana / rose" represent that the scent of banana is used to express fade-out in scents and the scent of rose is used to express fade-in in scents. We presented the scent which expresses fade-out before and the scent which expresses fade-in after, in spite of the intensity of the scents. Before conducting this experiment, we considered about all conceivable combinations of these four types of scents. When we used the scent of mint to express fade-out in scents ("mint / banana" etc.), we found that it was

Table 8: The number of changes of intensity in each combination of scents

| combination | Fade-out | Fade-in |
|---|---|---|
| banana / mint (Experiment 5.1) | 5 | 4 |
| banana / rose | 4 | 2 |
| banana / lavender | 4 | 2 |
| rose / banana | 4 | 3 |
| rose / mint | 5 | 4 |
| lavender / banana | 4 | 3 |
| lavender / mint | 5 | 4 |

difficult to feel the scents which express fade-in in scents because of refreshing feeling of mint. Also when we used the scent of lavender and rose ("lavender / rose" and "rose / lavender"), we found that it was difficult to distinguish these two scents because both of them were the scents of flower. Therefore, we omitted these combinations in this experiment.

The same as Experiment 5.1, we signaled the timing of breathing by sounds. After we presented 1 combination to the participant, we asked the participant how they felt. After taking a short break, we presented another 1 combination to the participant. We instructed them to answer by graphs, when they answered. Participants were 7 men and 5 women in their 20s.

### 5.3.2 Results and Consideration

In Experiment 5.1, we judged that participants could feel dissolve the most at "Dissolve 3", which they could feel five changes of the intensity in the scent of banana in fade-out and four changes of the intensity in the scent of mint in fade-in. Therefore, we examined the number of changes of the intensity of each scent and considered it comparing with the result of Experiment 5.1 that we showed in Section 5.2. Same as Section 5.2, we plotted the average values that are calculated from the values that each participant drew in each breath in each combination of the scents. The scent that expressed fade-out was presented in between the first breath and the sixth breath and the scent that expressed fade-in was presented in between the fourth breath and the ninth breath. That is, we presented two types of scents in between the fourth breath and the sixth breath. To examine whether the participants could feel the change of the intensity of scent between contiguous two breaths, such as the first breath and the second breath, the average values between contiguous two breaths for each scent were using t-test. If there was no significant difference between contiguous two breaths, the average values between n-th breath and (n+2)-th breath for each scent were analyzed. In this paper, we describe the result about "rose / mint" as an example. As a result of comparison in the scent of rose at "rose / mint", significant differences were found in between the first breath and the second breath, the second breath and the third breath, the third breath and the fourth breath and the fourth breath and the sixth breath ($p < 0.05$). The result of this indicates that the participants could feel changes of intensity of scent at the first, second, third, fourth and sixth breath. That is, the participants could feel five changes of the intensity in the scent of rose at "rose / mint". As a result of comparison in the scent of mint at "rose / mint", significant differences were found in between the fourth breath and the fifth breath, the fifth breath and the seventh breath and the seventh breath and the ninth breath ($p < 0.05$). The result of this indicates that the participants could feel changes of intensity of scent at the fourth, fifth, seventh and ninth breath. That is, the participants could feel four changes of the intensity in the scent of mint at "rose / mint". The similar comparison was also conducted for other combinations. Table 8 shows the number of changes of the intensity that the participants could feel in each combination of scents. Table 8 indicates that the number of changes of the intensity that participants could feel was the largest when the scent which expresses fade-in was mint. We obtained this result because the scent of mint was easy to distinguish from other three types of scents. In addition, this table indicates that it tends to be hard to feel the changes of the intensity of the scent when we choose two types of scents from the scent of banana, rose and lavender. Because the scents of banana, rose and lavender are classified in "sweet" scents, it is thought that it was not easy for participants to notice the changes of the intensity. Therefore, it is likely to feel dissolve in scents more when we use two scents which is easy to distinguish. Moreover, Osako et al. [15] studied that the unpleasant smell increases the sensory intensity. Therefore, it is thought that results of this experiment are possible to be changed when we present the unpleasant scents. In this experiment, we presented scents to participants many times, so we used the scent of banana, mint, lavender and rose which are relatively comfortable for human so as to reduce the discomfort of the participants.

## 6 CONCLUSION

Studies on transmitting scents together with various media to enhance the sense of reality are currently conducted. There are many scenes in videos and TV programs, such as the scene which many smelling objects appear at the same time, suddenly appear and gradually disappear. Therefore, it is necessary to control the presentation of scent in accordance with the changes in images/sounds over time. In doing so, it is more effective to enhance the sense of reality. However, the conventional presentation method of scents continues emitting scent at high density for a long time and gives problems of human adaptation to the scent and scents lingering. The receivers may not feel the intensity of scents properly. Therefore, it was difficult to present scents in accordance with videos.

To solve such problems, we studied about the change of kind of scents and the intensity of the scents by using pulse ejection which emits scent for very short periods of time. In this study, we paid attention to both the change of types and the intensity of scents and devised presentation techniques in scents changing the intensity of two types of scents.

Among these, we especially studied a presentation technique in scents that enables the receivers to feel “dissolve” which we defined as “the second scent becomes gradually strong at the same time as the first scent becomes gradually weak”. First, we developed presentation techniques in scents that the receivers to feel “fade-out” and “fade-in” using pulse ejection. As a result, the participants were likely to feel fade-out in scents and fade-in in scents the best in 6 phases. From these results, we combined 6 phases in fade-out and 6 phases in fade-in, and examined whether the participants can feel dissolve in scents. We aimed at the development of the presentation technique to express dissolve in scents with paying attention to the change of two scents and the intensity of the scents. The results of experiments revealed that participants could feel dissolve in scents the best by presenting fade-in and fade-out in scents which are overlapped in three breath. Based on this experiment, we examined whether the participant can feel dissolve in scents in different types of scents. As a result, it was revealed that it is likely to feel dissolve in scents when we use two scents which is easy to distinguish. It is expected that the technique can raise realistic sensations more when scents are presented in accordance with pictures by establishing the technique of dissolve in scents.

## REFERENCES


[1] Ministry of Posts and Telecommunications in Japan, Reports of the Association for Information and Communications Technology Using Five Senses (2007) (in Japanese).

[2] J. Naomi, G. Hunsaker, P.E. Gilbert and R.O. Hopkins, “The Role of Human Hippocampus in Odor-Place Associative Memory,” Chemical Sences, Vol.34, No.6, pp.513-521 (2009).

[3] C. Oshima, K. Nakayama and H. Ando, “Aroma That Enhances the Reality of Visual Images,” Review of the National Institute of Information and Communications Technology, Vol.56, Nos.1/2 (2010) (in Japanese).

[4] U. Haque, “Scents of space: an interactive smell system,” ACM SIGGRAPH 2004 Sketches (2004).

[5] ITmedia LifeStyle, the Lessons of Aromatherapy with Internet by OCN, http://www.itmedia.co.jp/lifestyle/articles/0505/23/news096.html (in Japanese).

[6] INTERNET Watch, NTT Com creates the movie called “New World” with a fragrance, http://internet.watch.impress.co.jp/cda/news/2006/04/11/11594.html (in Japanese).

[7] A. Kadowaki, J. Sato, Y. Bannai and K. Okada, “Presentation Technique of Scent to Avoid Olfactory Adaptation,” Proc.of ICAT 2007, pp.97-104 (2007).

[8] K. Ohtsu, J. Sato, Y. Bannai and K. Okada, “Measurement of Olfactory Characteristics for Two Types of Scent in a Single Breath,” IFIP INTERACT 2009, pp.306-318 (2009).

[9] D. Noguchi, K. Ohtsu, Y. Bannai and K. Okada, “Scent Presentation Expressing Two Smells of Different Intensity Simultaneously,” 2009 Joint Virtual Reality Conference of EGVE-ICAT-EuroVR, pp.53-60 (2009).

[10] S. Sugimoto, K. Ohtsu, Y. Bannai and K. Okada, “Scent Presentation Technique to Enable High Speed Switch of Scents,” The Journal of the Virtual Reality Society of Japan, Vol.15, No.SBR-1, pp.17-22 (2010) (in Japanese).

[11] J. Sato, K. Ohtsu, Y. Bannai and K. Okada, “Pulse Ejection Technique of Scent to Create Dynamic Perspective,” In proc. Of ICAT 2008, pp.167-174 (2008).

[12] CUERBO English Dictionary, http://english.cheerup.jp/eedict/search?name=dissolve

[13] Japan Association on Odor Environment, The guidebook of the Simplified measurement of Smell (2005).

[14] D. Noguchi, S. Sugimoto, Y. Bannai and K. Okada, “Time Characteristics of Olfaction in a Single Breath,” CHI’11 Proceedings of the 2011 annual conference on Human factors in computing systems, pp.83–92 (2011).

[15] M. Osano, K.Nishida and K. Shishida, “Impact of Odor Hedonics on Sensory Odor Intensity - Physiological intensity and sensory intensity-,” Environmental Health and Preventive Medicine (EHPM), Vol.44, No.5, pp.1002-1008 (1989).


(Received September 30, 2015)
(Revised December 12, 2015)

**Sayaka Matsumoto** B.S. degree from the Department of Information and Computer Science of Keio University, Japan in 2015. She is currently working toward an M.S. degree in Open and Environment Systems at Keio University. She received Best Paper Awards from IWIN 2015. Her research interests include fragrance information processing work.

**Shutaro Homma** B.S. degree from the Department of Information and Computer Science of Keio University, Japan in 2014. He is currently working toward an M.S. degree in Open and Environment Systems at Keio University. He received Excellent Presentation Awards from DICOMO 2015. His research interests include fragrance information processing work.

**Eri Matsuura** B.S. degree from the Department of Information and Computer Science of Keio University, Japan in 2014. She is currently working toward an M.S. degree in Open and Environment Systems at Keio University. Her research interests include fragrance information processing work.

**Shohei Horiguchi** B.S. degree from the Department of Information and Computer Science of Keio University, Japan in 2015. He is currently working toward an M.S. degree in Open and Environment Systems at Keio University. His research interests include fragrance information processing work.

**Ken-ichi Okada** received his B.S., M.S. and ph. D. in instrumentation engineering from Keio University, in 1973, 1975, and 1982, respectively. He is currently a chief professor in the Department of Information and Computer Science at Keio University. His research interests include CSCW, groupware, human computer interaction, and ubiquitous computing. He has published 140 journal papers, 170 international conference papers, and 18 books entitled "Collaboration and Communication," "Designing Communication and Collaboration Support Systems," "Introduction to Groupware" and so on. He is a member of IEEE, ACM, IEICE and IPSJ. He was a chair of SIGGW, a chief editor of IPSJ Magazine, and a chief editor of IPSJ Journal. Dr. Okada received the IPSJ Best Paper Award in 1995, 2000, and 2008, the IPSJ 40 the Anniversary Paper Award in 2000, IPSJ Fellow in 2002 and the IEEE SAINT Best Paper Award in 2004.

## Submission Guidance

### About IJIS

International Journal of Informatics Society (ISSN 1883-4566) is published in one volume of three issues a year. One should be a member of Informatics Society for the submission of the article at least. A submission article is reviewed at least two reviewer. The online version of the journal is available at the following site: http://www.infsoc.org.

### Aims and Scope of Informatics Society

The evolution of informatics heralds a new information society. It provides more convenience to our life. Informatics and technologies have been integrated by various fields. For example, mathematics, linguistics, logics, engineering, and new fields will join it. Especially, we are continuing to maintain an awareness of informatics and communication convergence. Informatics Society is the organization that tries to develop informatics and technologies with this convergence. International Journal of Informatics Society (IJIS) is the journal of Informatics Society.

Areas of interest include, but are not limited to:

Computer supported cooperative work and groupware
Intelligent transport system
Distributed Computing
Multi-media communication
Information systems
Mobile computing
Ubiquitous computing

### Instruction to Authors

For detailed instructions please refer to the Authors Corner on our Web site, http://www.infsoc.org/.

Submission of manuscripts: There is no limitation of page count as full papers, each of which will be subject to a full review process. An electronic, PDF-based submission of papers is mandatory. Download and use the LaTeX2e or Microsoft Word sample IJIS formats.

http://www.infsoc.org/IJIS-Format.pdf

LaTeX2e

LaTeX2e files (ZIP) http://www.infsoc.org/template_IJIS.zip

Microsoft Word™

Sample document http://www.infsoc.org/sample_IJIS.doc

Please send the PDF file of your paper to secretariat@infsoc.org with the following information:

Title, Author: Name (Affiliation), Name (Affiliation), Corresponding Author. Address, Tel, Fax, E-mail:

### Copyright



### Publisher

Address: Informatics Laboratory, 3-41 Tsujimachi, Kitaku, Nagoya 462-0032, Japan
E-mail: secretariat@infsoc.org

# CONTENTS