Who is interacting with me? Identification of an interacting person through playful interaction with a small robot

Reo Matsumura, Masahiro Shiomi, Takahiro Miyashita, Hiroshi Ishiguro, and Norihiro Hagita

***Abstract*—Small robots are being designed to recognize behaviors through playful interaction. Prior work used data from impoverished sensing devices such as inertial sensors to analyze gestures and attitude in playful interaction through time series analysis. However, the prior work did not focus on individual differences required for person identification. This research hypothesizes that person identification can be achieved by determining individual differences in playful interaction by using inertial sensor data. We propose a method that iteratively narrows down the candidates during interaction to achieve accurate person identification. This method calculates the features using a time series of the inertial sensor data. These features identify a candidate who is playfully interacting with the robot using a decision tree classifier that includes combinations of the current candidates. The system stores the results as a dataset for voting, and the voting results are used to reduce the candidates until the number of candidates is winnowed to one. Evaluation results show that our proposed method identifies persons through playful interactions with 99.1% accuracy.**

***Index Terms*—Identification of persons, Human-robot interaction, Humanoid robot**

I. INTRODUCTION

A. Inertial sensing through playful interaction

PLAYFUL interaction, which consists of sequences of intimate interaction patterns, is an essential mode of communication between people and small robots, such as pet-type and hobby robots. Researchers have found that playful interactions with robots are useful for therapy and enjoyment [1, 2]. Thanks to their small size, people can interact playfully with small robots as they would with a child or pet: picking them up, walking with them, and hugging them (Fig. 1).

However, this playful interaction complicates sensing. Due to social and cultural constructs, people identify themselves before engaging in other types of interaction, but such "polite" introductory behaviors cannot be expected with robots. Furthermore, since a robot is in very close proximity when held, it is difficult to use typical face recognition methods such as a camera on the robot (Fig. 1). Small robot technology would benefit from alternative identification methods that employ inexpensive sensors during close interactions.

In fact, the ability to sense the properties of the interacting person in playful interaction has become increasingly important in human-robot interaction. Researchers, by analyzing playful interaction, have developed such sensing mechanisms to recognize human gestures [3, 4], full-body gestures [5], and attitudes toward a robot [6]. Previous research simplified the processes of sensing and characterizing partners from playful interactions by using only time series data from inertial sensors.

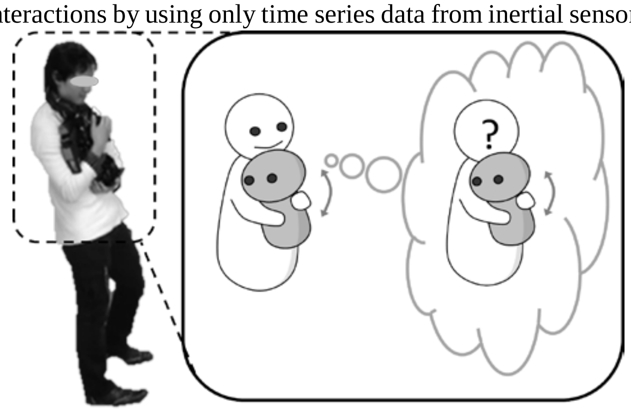

Fig. 1 Playful interaction with a small robot that cannot identify the interacting person, even when it can identify the interacting behavior

B. Person identification through playful interaction

Person identification through playful interaction with simple and inexpensive sensors (i.e., inertial sensors) offers great promise for small robots. For example, a person-identification capability would enable developers to design long-term interactions between robots and users and to personalize behaviors for users [7-9].

We hypothesize that person identification can be achieved by determining individual differences in playful interaction using inertial sensor data [2, 5]. For example, repeated playful behavior such as bouncing a baby greatly differs between parents. Past research that enabled small robots to identify interaction behaviors did not focus on person identification (details in Section II-A).

Manuscript received February 3, 2013. This research was supported by the Strategic Information and Communications R&D Promotion Programme (SCOPE), Ministry of Internal Affairs and Communications (132107010).

Reo Matsumura, Masahiro Shiomi, Takahiro Miyashita, and Norihiro Hagita are with ATR Intelligent Robotics and Communication Labs., Kyoto, Japan. Reo Matsumura is with Tokyo University, Tokyo, Japan. Hiroshi Ishiguro is with Osaka University, Osaka, Japan. (corresponding author's phone: +81-774-95-1432; fax: +81-774-95-1408; e-mail: m-shiomi@atr.jp).

C. Research aims

Here, we propose a method that uses inertial sensors to identify a person interacting with a small robot through playful interaction. We focus on inertial sensor data during each interaction pattern because observing sequences would increase the time needed for person identification. Our method applies two unique approaches: 1) We focused on the differences in extracted features from the inertial sensor data during playful interactions between individuals; 2) We iteratively extracted features from sensor output and narrowed down the candidates during interactions to identify the interacting person.

The rest of our paper is structured as follows. Section II describes related work, and Section III introduces our proposed method using inertial sensors to identify persons interacting through playful interaction. Section IV presents the experimental methods. Section V presents the results. Section VI provides a discussion, and Section VII summarizes the contributions.

II. Related Work

A. Recognition through interaction with small robots

Some research has included field trials in real environments with small animal robots [10] and small humanoid robots [11, 12] that playfully interact with people. These studies investigated how interactions with small robots affected children or the elderly, but they did not focus on person identification during such playful interaction.

Inertial sensors have been used for gesture and activity recognition in playful interactions with small robots. Using an accelerometer and tilt sensors, Salter et al. recognized four kinds of interaction between people and a ball-like robot: being alone, interaction, carrying, and spinning [3, 4]. MIT's Personal Robots group used a small teddy bear robot to recognize three gestures using inertial sensors: picking up, bouncing, and rocking [13, 14]. François et al.'s model classified interaction into two categories: gentle and strong [15]. Cooney et al.'s model recognized full-body gestures in playful interactions using inertial sensors and then designed enjoyable playful interaction with the model [2, 5]. This body of work clarified essential recognition functions for small robots through playful interactions, but it did not focus on person identification.

Information through physical interaction with environments might help in the identification process. For example, Giguere et al. developed a simple tactile probe used for accurate surface identification by Rumba, their mobile robot [16]. Miyashita et al. enabled small humanoid robots to efficiently recognize the environment by selecting their sensing behavior [17]. However, this prior work only focused on static targets and could not directly support person identification through playful interaction.

B. Person identification using inertial sensory information and other kinds of sensors

Researchers have developed person-identification systems by integrating inertial and other kinds of sensors. For example, Ikeda et al. developed a pedestrian-identification system by focusing on the correlation of rhythms calculated by wearable inertial sensors and a human tracking system [18]. Woodman et al. proposed a method to identify pedestrians using a Wi-Fi positioning method with inertial sensors [19].

As these efforts assumed the use of additional wearable sensors, they could not achieve person identification with inertial sensors alone; rather, they relied on a rich sensing environment such as a human tracking system or a Wi-Fi positioning system and information on the relative trajectory differences from other pedestrians. In contrast, we have realized person identification using only the robot's inertial sensors through playful interactions between people and small robots.

III. Identification of Interacting Person

A. Architecture

We proposed a method to identify a person who is engaged in playful interaction with a robot by phasing the reductions of candidates, i.e., a process of iteratively eliminating individuals from the pool of candidates. Fig. 2 shows an overview of our implemented system with inertial sensory information. We calculated the person identification features by a time series of the inertial sensor data (Section III-B). These features identify a candidate who is playfully interacting with the robot using a decision tree classifier that includes combinations of the current candidates (Section III-C). The system stores the results as a dataset for voting until the amount of data and the elapsed time from the last voting exceed thresholds for the next voting (Section III-D). The voting results are used to reduce the set of candidates (Section III-E). The system finally identifies the person who is playfully interacting with the robot by repeating these processes until the number of candidates is winnowed to one.

Note that our proposed method assumes that only one person is playfully interacting with the robot, data from the playful interaction are gathered beforehand, and the playful interactions continue for a certain period of time.

B. Feature extraction

Many kinds of features have been suggested by past research that used inertial sensors for various identifications [3–6, 16]. In this research, however, instead of using many features we employed a limited set of features to decrease the calculation costs, i.e., mean and standard deviation:

$$\mu = \frac{1}{n}\sum_{t-n}^{t} s(t) \qquad \textbf{(1),}$$

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{t-n}^{t}\left(s(t)-\mu\right)^2} \qquad \textbf{(2),}$$

where s(t) is the inertial sensor data at time t and n is the amount of data in the time period.

We also focused on the characteristics of playful interaction because they will probably be repeated and will enhance the identification of personal characteristics. In fact, past research reported that participants repeatedly interacted playfully with robots [2, 5]. Therefore, we also employed a zero-cross number, which is defined as the number of times the signal crosses the average value in time period n (Fig. 3). We used three kinds of features in each time step *f(T, μ, σ, z)* from each sensor to identify the person in playful interaction. Then these features were calculated by shifting one time step that equals the frequency of the sensor data (Fig. 4).

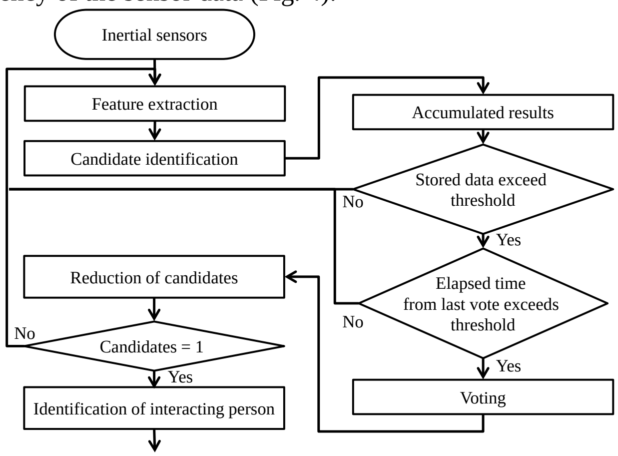

Fig. 2 Overview of proposed method

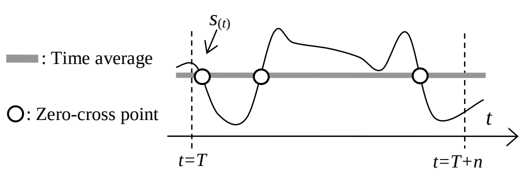

Fig. 3 Zero-cross number (a case of three zero-cross points)

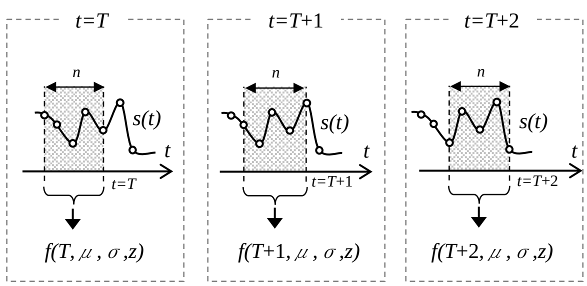

Fig. 4 Calculation of features using time-shift method

C. Candidate identification

In our proposed method, we applied a C4.5 decision tree classifier [20], which is widely used in robotics, to identify the correct candidate interacting with the robot at each time step. We prepared multiple decision trees that considered all combinations among candidates in advance of the phased reductions of the candidates during the voting process. Each decision tree includes at least two classes, and thus the number of decision trees is calculated by this function:

$$NumOfTree = 1 + \sum_{x=2}^{Nc-1} C(Nc, x) \qquad \textbf{(3),}$$

where *Nc* is the number of classes, i.e., the number of all candidates. In our proposed method, the system stores the results from a decision tree using *f(T, μ, σ, z)* for the voting process (details in next section). The candidate at each time step is calculated as

$$Candidate\ (t) = decision\,(RemainingCandidates,\ f\,(T, \mu, \sigma, z)) \qquad \textbf{(4),}$$

where *decision* is a function to output a candidate using a decision tree classifier that includes the *RemainingCandidates*. In the first process, *RemainingCandidates* includes all candidates: *RemainingCandidates* = {A, B … *Nc*}.

D. Voting process

Our proposed method's voting process determines the phased reductions of candidates using a set of candidate data within a certain time period (Fig. 5) instead of each candidate's result, since the system cannot know which timing is better to identify people during playful interactions. Since playful interactions are periodic [2,5], sometimes such interaction between individuals is quite similar or different within a single time period used for feature extraction. To avoid the difficulties of determining such timing, we used a voting process with a set of candidates from a certain time period; the effectiveness of this approach is investigated in the evaluation section.

The voting process is conducted every time period T_v by shifting a regular interval ($T_v/2$ sec) to avoid similar results between voting processes. In this process, the system calculates the selected ratio of each candidate in time period T_v:

$$Ratio\,(candidate) = \frac{1}{T_v}\sum_{x=0}^{T_v} count(if\,(candidate == Candidate(t-x))) \qquad \textbf{(5).}$$

E. Reduction of candidates

Next, the system narrows down the interacting person candidates based on their selection ratios. If a candidate's ratio is lower than a threshold, the candidate is removed in the next voting process. The threshold value is defined as

$$Th = 1/(NumOfCandidates * \alpha) \qquad \textbf{(6),}$$

where *NumOfCandidate* is the current number of candidates and α is a coefficient for the threshold calculation. The rest of the candidates are defined as

$$\begin{aligned} &RemainingCandidate = \\ &\{candidate \in CurrentCandidate \mid (Ratio(candidate) \geq Th)\} \end{aligned} \qquad \textbf{(7),}$$

where *CurrentCandidate* is the set of current candidates.

If *RemainingCandidate* includes more than two candidates, the system repeats the above processes until it has only one candidate.

IV. Experiment Methods

To investigate our proposed method's performance, we collected data where participants playfully interacted with a small robot equipped with inertial sensors.

A. Robot hardware

For the data collection, we used VisiON-4G, a small, human-like robot (Fig. 6) [21] that has a head (1 DOF), two arms (3 DOFs for each), a body, a waist, and two legs (7 DOFs for each). To record the inertial sensor data during the interactions, we attached two kinds of sensors to the robot's body: a 3-axis acceleration sensor (MMA7260Q, Freescale) and three 1-axis gyro sensors (ENC-03R, Murata Manufacturing) (Fig. 6). The characteristics of each sensor are shown in Tables 1 and 2. The sampling rate of each sensor was 20 Hz. Accordingly, the robot can record six axes of inertial information. We adopted a moving average method to reduce the noise of the sensor data.

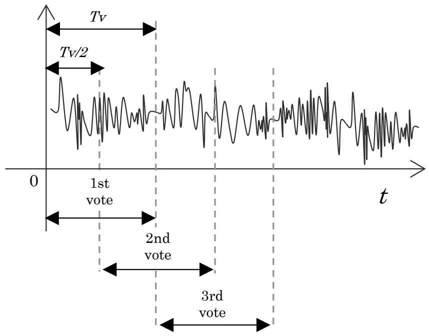

Fig. 5 Outline of voting process

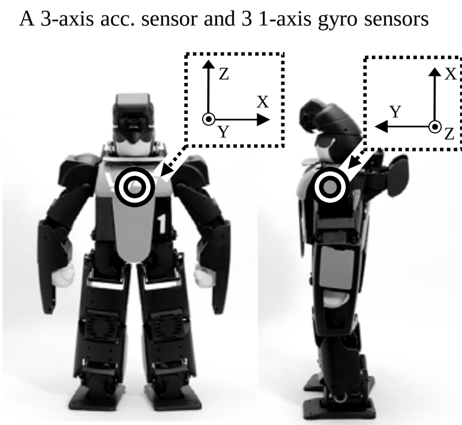

Fig. 6 Vision-4 G

Table 1 Specifications of 3-axis acceleration sensor

Model	MMA7260Q (Freescale)
Number of axes	3
Range	±6 g
Sensitivity	200 mV/g
Nonlinearity	±0.03%
Size	6.00 × 6.00 × 1.45 mm

Table 2 Specifications of 1-axis gyro sensor

Model	ENC-3R (Murata Manufacturing)
Number of axes	1
Range	±300 deg./sec.
Sensitivity	0.67 mV/deg./sec.
Nonlinearity	±5%
Size	4.0 × 8.0 × 2.0 mm

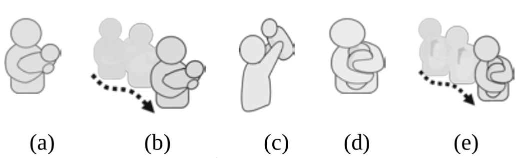

Fig. 7 Playful interaction patterns

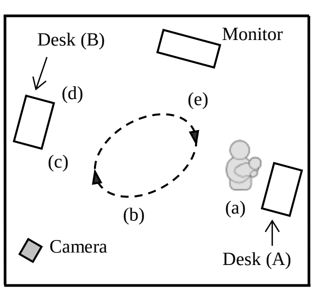

Fig. 8 Experiment environment

B. Participants

The participants were fifteen university students (eight men and seven women, average age: 20.8, S.D.: 1.7) recruited from the Internet without regard to major or specialty; their backgrounds were varied and most were not familiar with robots. They did not interact with our robot beforehand. They were paid 1,000 yen (roughly $12 U.S.) for one hour of participation.

C. Tasks

In the period of data collection, each participant playfully interacted with the robot based on the experimenter's instructions. In preparing to teach playful interactions, we referred to previous research that conducted playful interactions with small robots for data collection and chose five behaviors (Fig. 7): (a) horizontally hugging the robot like a child, (b) horizontally hugging it while walking, (c) bouncing it up and down as if amusing a child, (d) vertically hugging it, and (e) vertically hugging it while walking.

D. Procedure

We conducted our experiment in a laboratory (Fig. 8). First, the robot was placed on the right side of a desk (A). A monitor was placed in the back of the room. Before the first session, the participants were given a brief description of the experiment's purpose and procedure. Each participant conducted every interaction nine times. The interactions were displayed on a monitor. The order of the instructed interactions was counterbalanced, and each participant interacted with the robot 45 times (Fig. 9). The experimenter asked the participants to walk around the center of the room or between the desks during the walking behaviors.

Each interaction was conducted within 51.2 seconds due to the limitations of the recording system. Thus, the image displayed on the monitor changed with a beep every 51.2 seconds. The first 6.4 seconds of data were not used to avoid the effects of interaction transitions.


Fig. 9 Data collection as a participant hugged the robot and walked

E. Hypothesis and prediction

For person identification using small robots equipped with inertial sensors, we focused on personal characteristics in the interaction patterns of playful interactions. Since such personal characteristics are different among individuals, we assume that a robot can identify the interacting person if it can distinguish such differences through interaction. Personal characteristics such as habits and rhythm probably appear through repetition of an interaction pattern in playful interaction during a certain time period. Therefore, it is difficult to use a single moment of sensor data, or a summary of all sensor data during interactions, to identify differences between personal characteristics from the viewpoint of sensor output.

Consequently, we applied voting to reduce the candidates rather than allowing an increase in the *needed data length*, which is the number of data points from the time series that are processed by an algorithm before the algorithm terminates with a classification. Our approach is useful for observing the amount of difference between personal characteristics. We also assume that such a step-by-step sensing process will be more effective than a process that summarizes all of the sensor data, because summarizing complicates the task of finding differences in personal characteristics and needs much calculation time.

Our hypothesis states that if we successfully implement our ideas, our proposed method will accurately identify the interacting person through playful interaction. We predict the following:

Prediction 1: Our proposed method, which considers personal characteristics based on inertial sensor data and voting processes, will achieve a higher *identification ratio* than an alternative method that summarizes all sensor data during interactions for person identification. Moreover, our proposed method will decrease the *needed data length* from the level of the alternative method.

Prediction 2: The proposed method with voting processes will achieve a higher *identification ratio* than the proposed method without voting processes. On the other hand, using the voting processes will increase the *needed data length*.

F. Dependent measures

We measured the *identification ratio*, which is considered the success rate of the system correctly identifying the interacting person. We also measured the *needed data length*, which is the length of the data used for person identification.

G. Experimental Design and Data Analysis

15 participants each interacted 45 times with the robot. We conducted a leave-one-out cross-validation to measure the performance of the person identification. We used 674 datasets to construct decision tree classifiers and one dataset to test the identification in each validation. The system used 674 datasets to generate 32,752 trees (sum of ${}_{15}C_i$ (i=2~14)). Note that these processes can be done before the experiment but not in real time.

We set each parameter to construct the decision tree classifiers: n = 64 steps (3.2 sec), and T_v = 128-step (6.4 sec) intervals of voting time, i.e., T_v /2 = 64 steps (3.2 sec). We determined the coefficient of α in Eq. (6) to be four. These parameters are based on heuristic tuning with the test data.

In this experiment, we prepared two alternative methods to investigate the effectiveness of our proposed method: *proposed method without voting process* and *simple identification method*. The former investigates the effectiveness of the voting process in our proposed method; the latter investigates whether our approach is effective by focusing on the personal characteristics in the time series data of the inertial sensors.

The *proposed method without voting process* reduces the candidates using the candidate results at each time step instead of the voting process; the right part of Fig. 2 is not used for person identification. For this purpose, this method uses the reliability from a confusion matrix of a decision tree classifier. For example, for the confusion matrix of the decision tree classifier shown in Table 3, the reliability is 80% when the decision tree classifier's results are "A" and 60% when its results are "B." Accordingly, the following equation is used instead of (5). We used the same threshold value of the proposed method, i.e., (6), and the coefficient of α is four.

Table 3 Example of a confusion matrix

a	b	
80	20	a = Class A
40	60	b = Class B

$$Ratio(t) = ConfusionMatrix\,(Remainingcandidate, candidate\,(t)) \quad \textbf{(8).}$$

The *simple identification method* did not use the voting process or the phased reductions of the candidates. It simply stored all of the sensor data, calculated the features from them (i.e., n = 896 steps = 44.8 sec), and used a decision tree classifier only once to calculate the candidates. Thus, the number of decision tree classifiers is one in this method, which includes all candidates.

Fig. 10 Vertically hugging robot while walking: Participant A

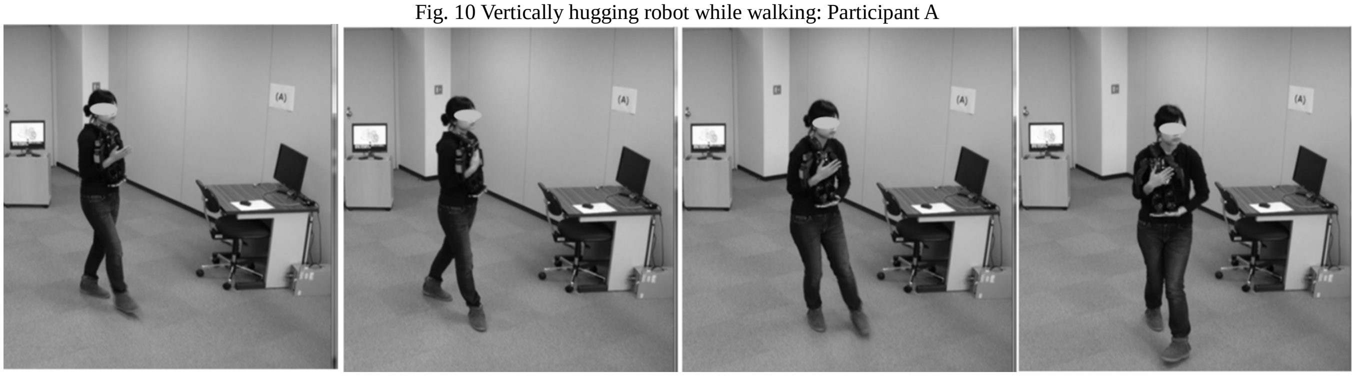

Fig. 11 Vertically hugging robot while walking: Participant B

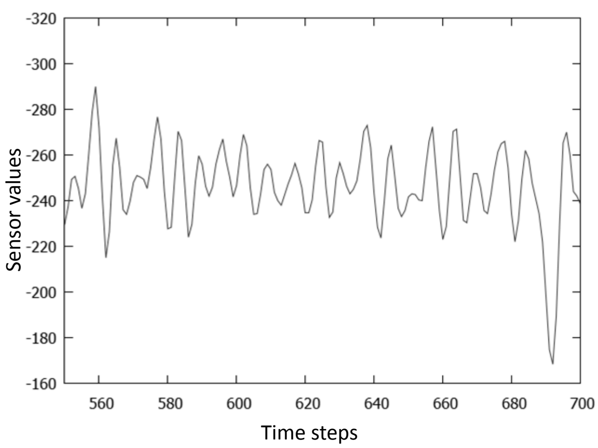

Fig. 12 Sensor data of Z axis of acc. sensor from Participant A

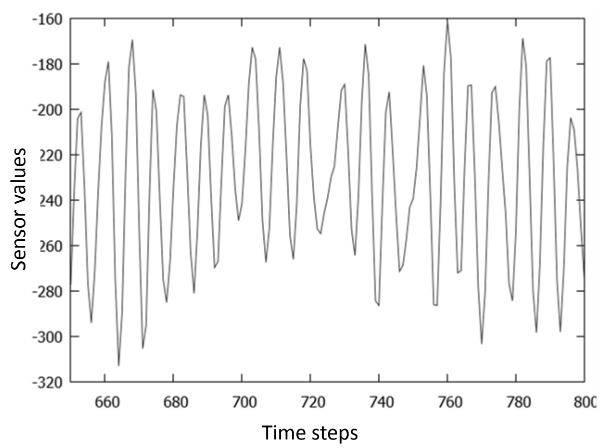

Fig. 13 Sensor data of Z axis of acc. sensor from Participant B

V. Results

A. Evaluation of person identification

1) Verification of predictions

Table 4 shows the results of the person identification and the calculation time for each method. Fig. 14 shows a confusion matrix with the proposed method, which achieved 99.1% for person identification through playful interactions; the alternative methods only achieved 92.4% and 72.4% accuracy.

Cochran's Q test revealed significant differences in the identification rate between conditions (Q = 243.750, p<.001). Multiple comparisons also revealed significant differences: *proposed* > *proposed method without voting process* (p=.001), *proposed* > *simple identification* (p<.001), and *proposed method without voting process* > *simple identification* (p<.001).

We also conducted a repeated measures ANOVA and found a significant main effect (p<0.001, partial η2 = 1.00). A multiple comparison by the Bonferroni method revealed that the *needed data length* of the *proposed* condition was significantly less than for *simple identification* (p. <001) and significantly more than for *proposed method without voting process* (p<.001). We used this analysis method because the datasets were the same between conditions, which were prepared by the leave-one-out cross validation method.

Our results show the effectiveness of the voting process using phased reductions of candidates for person identification. Moreover, the *needed data length* of the proposed method with the voting process was lower than that of the alternative method but still higher than that of the proposed method without the voting process. Therefore, our predictions were supported.

Table 4 Evaluation results

Type of method	Identification rate		Needed data length (seconds)
	Average (%)	S.D.	
Proposed	99.1	1.6	16.1
Proposed method without voting process	92.4	6.1	9.7
Simple identification	72.4	7.5	44.8

	A	B	C	D	E	F	G	H	I	J	K	L	M	M	O
A	43	0	0	0	0	2	0	0	0	0	0	0	0	0	0
B	0	44	0	0	0	0	0	0	0	0	0	1	0	0	0
C	2	0	43	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	44	0	1	0	0	0	0	0	0	0	0	0
E	0	0	0	0	45	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	45	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	45	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	45	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	45	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	45	0	0	0	0	0
K	0	0	0	0	0	0	0	0	0	0	45	0	0	0	0
L	0	0	0	0	0	0	0	0	0	0	0	45	0	0	0
M	0	0	0	0	0	0	0	0	0	0	0	0	45	0	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	45	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45

Fig. 14 Confusion matrix with proposed method

2) Interaction differences among participants

Our hypothesis has already been supported by comparing the identification rates, but showing the differences in interaction styles is also important to verify its validity. Therefore, we present sensor data and images that show how playful interactions with the robot differed among the participants.

Figs. 10 and 11 show playful interaction scenes between the robot and two participants. Participant A held the robot in both her arms and frequently bent over while walking. Therefore, the robot was held firmly to her body by the pressure from her arms, but its body swings with her. Participant B put her left hand under the robot's foot and her right hand on its back (Fig. 11). The robot's body moved up and down with her walking behavior. She also frequently patted its back while walking.

These observations suggest that such personal characteristics as mannerisms appear through playful interactions. In fact, these different interaction styles exhibit different trends of sensor data (Figs. 12 and 13); we see that these trends are quite different in this dimension. Such differences are essential for identifying interacting people with inertia sensors. We believe that these findings support our hypothesis and the validity of our approach.

3) Analysis of individual differences among participants

We analyze how well our proposed method identifies participants through the phased reductions of candidates. If the method works properly, the sensor data from the inertial sensors can categorize each participant through the voting processes. For this analysis, we tried to visualize the sensor features that are calculated by different sensors and that have different units. Due to the difficulty of comparing such bits of data, we used features that have the highest information gain ratios in the calculations depending on the analysis setting (i.e., combinations of candidates). We adopted the multi-dimensional scaling (MDS) method with the classical multi-dimensional scaling of a data matrix, known as the principal coordinates analysis method [22], to visualize the relationships among each participant's playful interactions.

Fig. 15 shows the relative relationships among all nine interactions of "horizontally hugging robot while walking" calculated by MDS. It used the mean values from the acceleration sensor on the x, y, and z axes that are most often used in decision tree classifiers. Part of the relative distance among participant is large enough, e.g., participants D, F, and I. Some participant results showed large variance in their own interactions, e.g., participant B, but they still have enough distance from the others. On the other hand, this result indicates the difficulties of identifying an interacting participant with a single classification process because the relative distances among some participants are similar: participants E, H, and K.

Fig. 16 also shows the relative relationships after the phased reductions from Fig. 14; it only includes the nine interactions of participants E, H, and K. Due to the differences among the features of a decision tree classifier that includes all candidates and one that includes only three participants , we used the SDs of the z axis on the acceleration and gyro sensors and the zero-cross number of the y axis on the gyro sensor.

The relative distances among participant K and the others are larger than those values for the other participants. Note that the next voting process could identify the interaction participant using a decision tree that includes participants E and H. We believe that this observation supports the ability of our proposed method with a voting process to effectively phase the reductions of candidates by focusing on personal characteristics at the level of sensor data, even if parts of the playful interaction style are similar among candidates.

B. Evaluation of behavior recognition

We discuss the behavior recognition performance of our proposed method through playful interaction, in relation to that achieved by past research [2]. If our proposed method can be applied to such recognition in the same manner and show high performance, it will also be useful for the development of small robots.

For person identification, we again conducted leave-one-out cross-validation to measure the behavior recognition performance. In all, 674 datasets were used to construct decision tree classifiers, and one dataset was used to test the identification in each validation with the same parameters. In this evaluation, first, the system identifies the participant and then uses the same test data to identify the category of playful interaction. We defined a case as successful when both the participant and the category are correctly identified; if either the participant or the category is incorrect, it is defined as a failure.

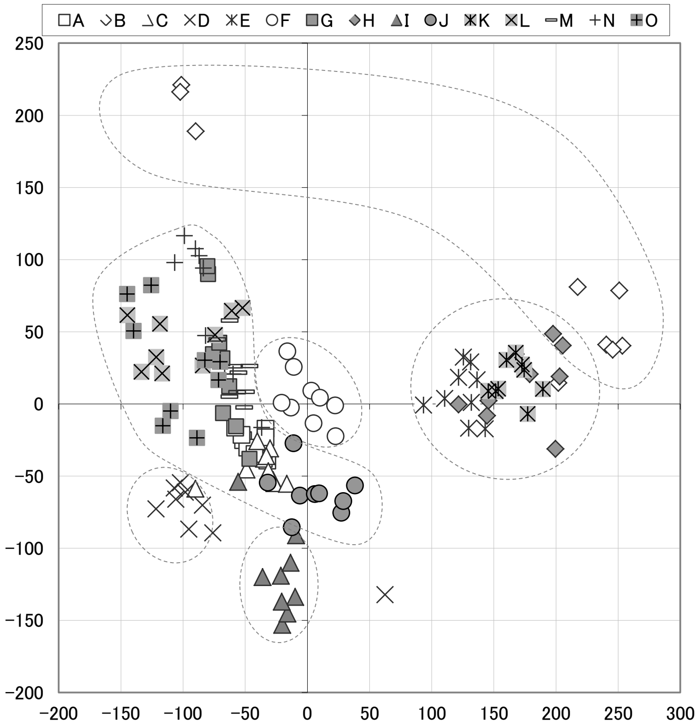

Fig. 15 MDS features of all participants' features

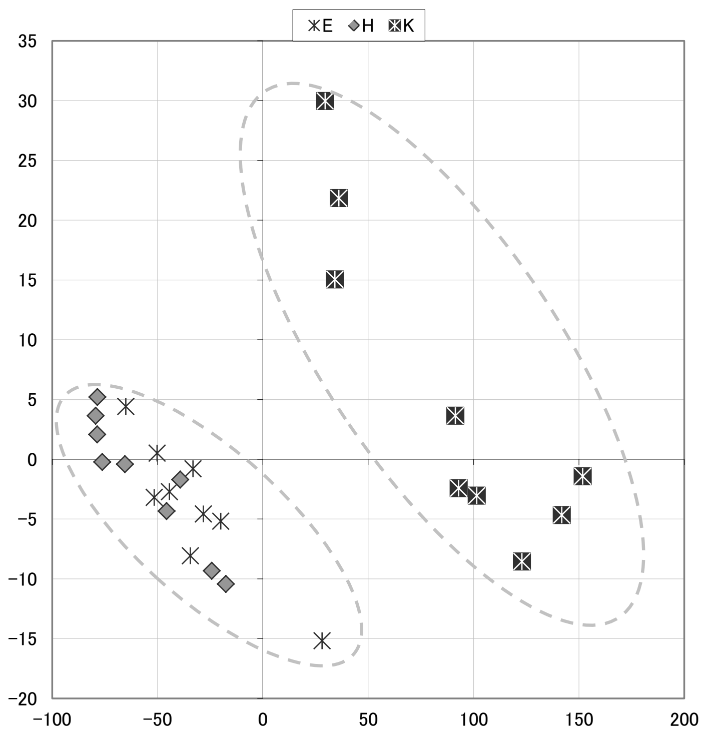

Fig. 16 MDS features of E, H, and K's features after a voting process

Table 5 shows the person identification results for each method. Our proposed method achieved 80.6% accuracy for person and behavior identification through playful interactions. The alternative methods only achieved 65.9% and 50.1% accuracy.

Cochran's Q test revealed significant differences in the identification rate among conditions (Q = 162.880, p<.001). Multiple comparisons also revealed significant differences: *proposed* > *proposed method without voting process* (p<.001), *proposed* > *simple identification* (p<.001), and *proposed method without voting process* > *simple identification* (p<.001). Our results again show the effectiveness of the voting process for the phased reductions of candidates for behavior recognition, as with the person identification results.

Table 5 Evaluation results of behavior recognition

Method	Identification rate		Needed data length (seconds)
	Average (%)	S.D.	
Proposed	80.6	8.9	17.0
Proposed method without voting process	65.9	8.9	9.7
Simple identification	50.1	6.5	44.8

VI. Discussion

A. Scalability

1) Number of candidates

We evaluated the proposed method with data from 15 participants using 18 sensor features. If we increase the number of targets, our proposed method should identify them, although it might require more interaction time. As shown in Figs. 15 and 16, the method identified the interacting participants through contentious voting processes, even when the first voting process failed to clearly separate multiple participants. With more targets, similar interaction patterns will also increase, so that more interaction time is needed for identification.

In fact, the calculation time of the proposed method becomes longer due to the voting process shown in Tables 4 and 5. From the viewpoint of playful interaction, we do not expect these extended times to be a problem, since we can reduce the calculation time by adjusting the voting threshold; however, this solution involves a performance trade-off. Accurate identification is more important than relatively faster calculation time because it involves such essential information as personal features.

2) Time universality

We used the time series of the sensor data for person identification. If the difference in interaction patterns were substantially different due to time changes (e.g., morning and evening), using our proposed method as a person identification method would be difficult. We address this problem by analyzing the first and last bits of interaction data from the data collection. Since this time period includes at least 55 minutes, a rest time, and other kinds of interactions, it may be adequate for discussing individual changes caused by the progression of time.

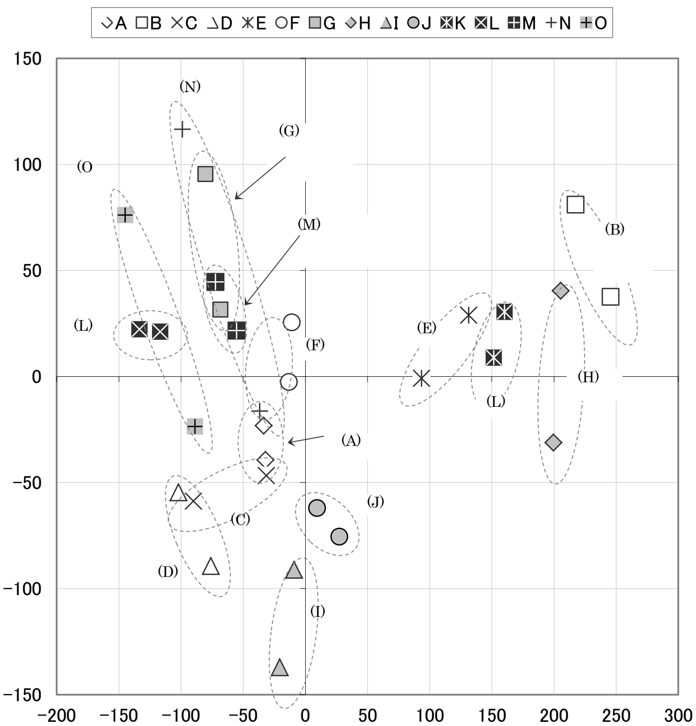

Fig. 17 MDS features on 1st & 9th trials of all participants

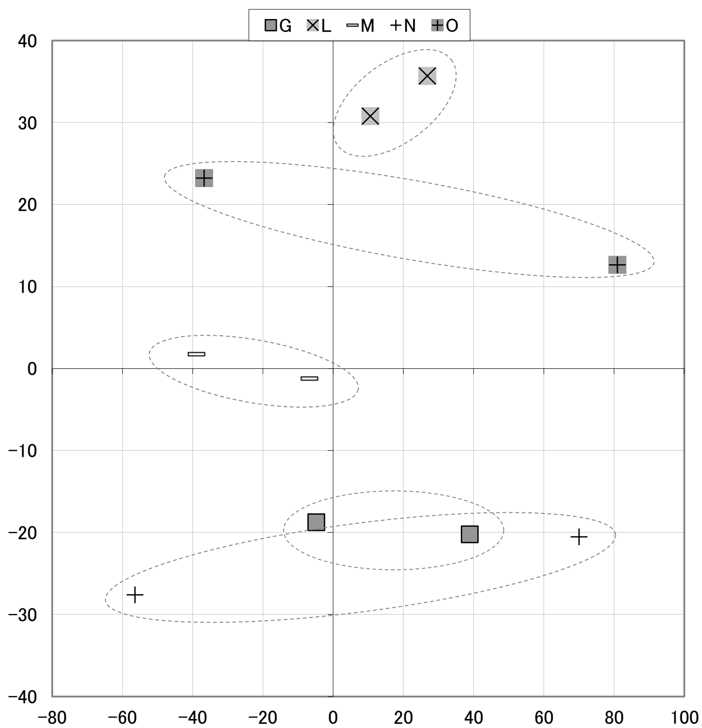

Fig. 18 MDS features on 1st & 9th trials of G, L, M, N, and O

Fig. 17 shows the relationship between each first and last interaction of “horizontally hugging robot while walking” calculated by MDS. The features are the mean values on the x, y, and z axes of the acceleration sensor. This figure shows a similar trend to Fig. 15; part of the dataset can be separated to different clusters, but separating the other data is difficult. Fig. 18 also shows the relative relationships after phasing the reductions from Fig. 17. Therefore, this figure only includes the first and last interactions of persons G, L, M, N, and O. For this calculation, since we used the x, y and z axes of the standard deviation of the acceleration sensor, this figure also shows a similar trend to that of Fig. 16, indicating that the differences between the first and last interactions of the same person are slightly smaller than those of other people’s interaction patterns. Consequently, we do not believe the time intervals significantly affect people’s interaction patterns.

3) *Robot appearance*

The interaction style humans assume with a robot may be influenced by its appearance. In this experiment, the robot had a human-like appearance and was about the same size as a baby. Therefore, playful interaction with it can be designed in a similar way as interaction with a baby. If the robot’s appearance were different from a human-like appearance (e.g., a ball, a creature, or a toy), the playful interaction style may be different than with human-like robots

Such differences based on robot appearance are one limitation of this work. In this experiment, we only confirmed the effectiveness of our proposed method with a human-like robot. However, even when the robot’s appearance is different from that of a human-like robot, playful interaction with a robot may be different among individuals because it is influenced not only by personality but also by such physical properties as strength. Therefore, we believe that our proposed method can be adopted for robots with non-humanlike appearances.

B. *Parameter tuning*

We set the parameters for feature extraction based on heuristic tuning. This is another limitation of this research. In this paper, we heuristically found the parameters that provide adequate performance for person identification. However, parameters would need to be re-tuned for different settings, such as the number of interacting people, different personal characteristics like the differences between adults and children, and the shape/weight of the robot. In fact, if the robot is too heavy, interaction styles may be different due to personal strength. If the robot's appearance is not human-like, participants may also choose different interaction styles.

To automate parameter tuning, such basic ways as a grid search and a generic algorithm can be applied. These methods can also use parameter tuning for generating decision trees. Automation of parameter tuning remains as future work.

C. *Applications and future work for long-term interaction*

We investigated whether about one hour (55 minutes) of repetition would be a problem for person identification. However, it is important to consider how the system updates stored data for person identification for such long-term situations as over one month. Moreover, the transition of interacting persons would be a problem for person identification. In the real world, we need to consider such transitions.

To solve these types of problems, we might need at least three kinds of implementations. The first is storing interaction data and continuously updating decision trees in a background process to deal with the changes in interaction style due to long-term interaction. The second is dealing with people transitions. For this purpose, the system needs to continuously calculate the reliability of the current interacting person even if the identification is finished. This function is also needed to update decision trees for long-term interaction. The final one is integration with other identification methods, such as face recognition with environmental cameras. Our proposed method achieved high performance with only inertia sensors, which can cover situations where a robot closely interacts with a person. By integrating our proposed method with existing systems, we might develop more accurate person identification systems under several situations, particularly in real environments. In this paper, we focused on developing a novel approach for person identification, not the integration of technologies; however, for future work, we will integrate our method with existing methods.

D. Interaction from robot for person identification

We did not evaluate the effects of interaction from the robot during playful interaction. Future work should consider such effects for person identification, because people who are interacting with the robot will change their behavior when it is talking or moving during interactions. The reaction of people toward interaction from a robot would include personal characteristics and differences with others; therefore, it would also be useful for person identification.

E. Limitations

We acknowledge other limitations. Our proposed method identified persons through playful interactions with 99.1% accuracy, but we did not compare its performances with other state-of-the-art time series classification algorithms, such as Support Vector Machine [23], an existing method that uses inertia sensors to identify environment [16], time series shapelets [24], and time series data mining [25]. Using these methods might increase performance and decrease the needed data length.

This method can identify only one person because it assumes that a small robot interacts with only one person at a time. Moreover, winnowing candidates requires observation of a playful interaction within a certain time to find differences between individuals. Thus, if the length of interaction is shorter than the average time of needed data length, the performance of person identification will decrease. In other words, the method assumes a continuous playful interaction within a certain time period that is a relatively longer than the times of other types of person identification systems.

Identification is conducted using an existing data set. Therefore, a registration process such as an interaction with a robot is needed beforehand. Such processes are also needed for other types of person identification such as face/fingerprint-based systems, but the proposed method needs more time for registration than these person-identification methods.

We tested only adult participants with specific playful interaction patterns. Therefore, we might need to adjust the parameters or the sensor features to adapt our proposed method to other types of participants, such as children or senior citizens, and different kinds of interaction patterns in real-life settings.

We evaluated our method in an experimental scenario, but it would be interesting to test whether other person identification methods, such as facial recognition, can identify interacting persons in the same scenario. This is also important for future integration between our proposed method and other methods, as discussed in Section VI-C.

VII. Conclusion

In this paper, we presented a novel method that identifies persons through playful interaction with a small robot. The unique concept of this work is its focus on the differences in extracted features from the inertial sensor data during playful interaction to identify the interacting person. This approach is different from related work that focused on eliminating individual differences to identify the interacting behavior. Our proposed method iteratively extracts features from the inertial sensor data history and narrows down the interacting person candidates during interaction to identify the person. We experimentally evaluated the performance of our proposed method, and our evaluation results showed that it identified persons through playful interactions with 99.1% accuracy. This work might be useful with various small robots, such as hobby-type robots. Person identification is an essential function for interacting with people, but it is physically difficult during close interactions. Our proposed method enables such robots to identify interacting persons during playful interactions.

References

[1] K. Wada and T. Shibata, "Living With Seal Robots-Its Sociopsychological and Physiological Influences on the Elderly at a Care House," *IEEE Transactions on Robotic*s, vol. 23, No. 5, pp. 972-980, 2007.

[2] M. D. Cooney, T. Kanda, A. Alissandrakis, and H. Ishiguro, "Interaction Design for an Enjoyable Play Interaction with a Small Humanoid Robot," *Humanoids 2011*, pp. 112-- 119, 2011.

[3] T. Salter, F. Michaud, K. Dautenhahn, D. Létourneau, and S. Caron, "Recognizing interaction from a robot's perspective," *Proceedings IEEE International Workshop on Robot and Human Interactive Communication*, pp. 178-183, 2005

[4] T. Salter, F. Michaud, D. Letourneau, D. C. Lee, and I. P. Werry, "Using Proprioceptive Sensors for Categorizing. Human-Robot Interactions," *2nd ACM/IEEE International. Conference on Human-Robot Interaction*, 2007.

[5] M. D. Cooney, C. Becker-Asano, T. Kanda, A. Alissandrakis, and H. Ishiguro, "Full-body Gesture Recognition Using Inertial Sensors for Playful Interaction with Small Humanoid Robot," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2010)*, 2010.

[6] D. François, D. Polani, and K. Dautenhahn, "On-line behaviour classification and adaptation to human-robot interaction styles," *Proceedings of the ACMIEEE International Conference on Humanrobot Interaction,* pp. 295-302, 2007.

[7] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro," Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial," *Human Computer Interaction (Special issue on human-robot interaction),* Vol. 19, No. 1-2, pp. 61-84, 2004.

[8] T. Kanda, M. Shiomi, Zenta Miyashita, H. Ishiguro, and N. Hagita, A communication robot in a shopping mall, *IEEE Transactions on Robotics*, Vol. 26, No. 5, pp. 897-913, 2010.

[9] A. M. Sabelli, T. Kanda, and N. Hagita, "A Conversational Robot in an Elderly Care Center: an Ethnographic Study," *6th ACM/IEEE International Conference on Human-Robot Interaction (HRI2011)*, pp. 37-44, 2011.

[10] T. Shibata and K. Tanie, "Physical and Affective Interaction between Human and Mental Commit Robot," *IEEE Int. Conf. on Robotics and Automation (ICRA2001)*, pp. 2572-2577, 2001.

[11] F. Tanaka, A. Cicourel, and J. R. Movellan, "Socialization between Toddlers and Robots at an Early Childhood Education Center," *Proceedings of the National Academy of Sciences of the USA (PNAS)*, pp. 17954-17958, 2007.

[12] F. Tanaka, J. R. Movellan, B. Fortenberry, and K. Aisaka, "Daily HRI Evaluation at a Classroom Environment: Reports from Dance Interaction Experiments," *ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI2006)*, pp. 3-9, 2006

[13] J. K. Lee, R. L. Toscano, W. D. Stiehl, and C. Breazeal, "The Design of a Semi-Autonomous Robot Avatar for Family Communication and Education," *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2008.

[14] W. D. Stiehl, J. Lieberman, C. Breazeal, L. Basel, L. Lalla, and M. Wolf, "Design of a Therapeutic Robotic Companion for Relational, Affective Touch," *IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN)*, 2005.

[15] D. Francois, D. Polani, and K. Dautenhahn, "Online Behaviour Classification and Adaptation to HumanRobot Interaction Styles," *IEEE-RAS International Conference on Humanoid Robots*, pp. 353-359, 2008.

[16] P. Giguère and G. Dudek, "A Simple Tactile Probe for Surface Identification by Mobile Robots," *IEEE Transactions on Robotics* Vol. 27, No. 3, pp. 534-544, 2011.

[17] T. Miyashita, K. Shinozawa, N. Hagita, and H. Ishiguro, "Behavior Selection and Environment Recognition Methods for Humanoids based on Sensor History," *IROS 2006*, pp. 3468-3473, 2006

[18] T. Ikeda, H. Ishiguro, T. Miyashita, and N. Hagita, "Pedestrian identification by associating walking rhythms from wearable acceleration sensors and biped tracking results," *in Proc. of PECCS*, 2012.

[19] O. Woodman and R. Harle, "Pedestrian Localization for Indoor Environment," *In Proc. UbiComp 2008*, pp. 114-123, 2008

[20] J. R. Quinlan, "C4.5 Programs for Machine Learning," Morgan Kaufman, 1993.

[21] R. Matsumura and H. Ishiguro, "Development of a High-Performance Humanoid Soccer Robot," *International Journal of Humanoid Robotics*, Vol. 5, No. 3 (2008) 353-373.

[22] J. C. Gower, "Some distance properties of latent root and vector methods used in multivariate analysis," Biometrika vol. 53, 325–328, 1966.

[23] V. Vapnik.: "The Nature of Statistical Learning Theory," Springer, (1995).

[24] L. Ye and E. J. Keogh, "Time series shapelets: a novel technique that allows accurate, interpretable and fast classification," Data Min. Knowl. Discov., vol. 22, no. 1-2, pp. 149–182, 2011.

[25] T. Chung Fu, "A review on time series data mining," Eng. Appl. of AI, vol. 24, no. 1, pp. 164–181, 2011

Reo Matsumura received M.Eng. and Ph.D. degrees in engineering from Osaka University in 2008 and 2012. From 2008 to 2012, he was an intern researcher at the Intelligent Robotics and Communication Laboratories (IRC) and is currently a cooperative researcher at IRC, a project researcher at Tokyo University, and an engineer at YUKAI-Engineering Co., Ltd. in Tokyo. His research interests include the development of communication robots and their control theory.

Masahiro Shiomi received M.Eng. and Ph.D. degrees in engineering from Osaka University in 2004 and 2007. From 2004 to 2007, he was an intern researcher at the Intelligent Robotics and Communication Laboratories (IRC) and is currently a group leader at Agent Interaction Design department, IRC at the Advanced Telecommunications Research Institute International (ATR) in Kyoto. His research interests include human-robot interaction, interactive humanoid robots, networked robots, and field trials.

Takahiro Miyashita received B.S., M.S., and Ph.D. degrees in engineering for computer-controlled machinery from Osaka University in 1993, 1995, and 2002. From 1998 to 2000, he was a fellow researcher of the Japan Society for the Promotion of Science. From April to September 2000, he was a researcher at the Symbiotic Intelligent Group at the ERATO Kitano Symbiotic Systems Project of Japan Science and Technology Agency. From October 2000 to July 2002, he was an assistant professor at Wakayama University. Currently, he is a group leader and a senior research scientist at the Intelligent Robotics and Communication Laboratories (IRC), Advanced Telecommunications Research Institute International (ATR). He is also a member of the Robotics Society of Japan, the Japan Society of Mechanical Engineers, and the Japanese Society for Artificial Intelligence.

Hiroshi Ishiguro received a D.Eng. in systems engineering from Osaka University in 1991. From 2009, he has been a professor in the Department of Systems Innovation of the Graduate School of Engineering Science at Osaka University. From 2002, he has been a group leader of the Hiroshi Ishiguro Laboratory at the Advanced Telecommunications Research Institute and an ATR fellow. He was a research associate from 1992 to 1994 in the Graduate School of Engineering Science at Osaka University and an associate professor from 1998 to 2000 in the Department of Social Informatics at Kyoto University. From 1998 to 1999, he was a visiting scholar at the University of California, San Diego. He was an associate professor from 2000 to 2001 and a professor from 2001 to 2002 in the Department of Computer and Communication Sciences at Wakayama University. From 2002 until 2009, he was a professor in the Department of Adaptive Machine Systems in the Graduate School of Engineering at Osaka University. His research interests include distributed sensor systems, interactive robotics, and android science. His android studies are very well known in the field. He received the Osaka Cultural Award in 2011.

Norihiro Hagita received B.S., M.S., and Ph.D. degrees in electrical engineering from Keio University, Tsuruoka City, Japan in 1976, 1978, and 1986. From 1978 to 2001, he was with the Nippon Telegraph and Telephone Corporation (NTT). He joined the Advanced Telecommunications Research Institute International (ATR) in 2001 to establish the ATR Media Information Science Laboratories and the ATR Intelligent Robotics and Communication Laboratories in 2002. His current research interests include communication robots, networked robot systems, interaction media, and pattern recognition. Dr. Hagita is a fellow of the Institute of Electronics, Information, and Communication Engineers, Japan. He is a member of the Robotics Society of Japan, the Information Processing Society of Japan, and The Japanese Society for Artificial Intelligence. He is a co-chair for the IEEE Technical Committee on Networked Robots.