The final publication is available at
http://www.tandfonline.com/doi/full/10.1080/01691864.2015.1095651

Finding a Person with a Wearable Acceleration Sensor using a 3D Position Tracking System in Daily Environments

Masahiro Shiomi and Norihiro Hagita

ATR-IRC, Kyoto, Japan

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). He is currently a group leader in the Agent Interaction Design department at IRC, Advanced Telecommunications Research Institute International (ATR). His research interests include human-robot interaction, robotics for child-care, networked robots, and field trials.

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 and established 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. He is a fellow of the Institute of Electronics, Information, and Communication Engineers, Japan as well as a member of the Robotics Society of Japan, the Information Processing Society of Japan, and The Japanese Society for Artificial Intelligence. He is also a co-chair for the IEEE Technical Committee on Networked Robots.

The final publication is available at
http://www.tandfonline.com/doi/full/10.1080/01691864.2015.1095651

Finding a Person with a Wearable Acceleration Sensor using a 3D Position Tracking System in Daily Environments

Person identification with accurate position information is essential for providing location-based services in real environments, such as a shopping mall. For this purpose, we propose a method that integrates 3D position information from environmental depth sensors and acceleration data from wearable devices to anonymously gather the trajectories of people who have wearable devices as well as others. Our proposed method identifies a person who has a wearable device by comparing two time-series of acceleration data from device and position information. To do this, we extracted the behaviours of each axis using the changes of each bit of acceleration data at certain time periods. We evaluated our method with data collected at a shopping mall and a children's playroom to investigate its effectiveness and robustness in different environments. Our evaluation results showed that it achieved an average identification of 85%, which is better than several alternative methods.

Keywords: person identification; acceleration sensor

1. Introduction

Acceleration information is one common type of information for developing sensing systems in the research field of robotics. Since acceleration sensors are inexpensive, tiny, and easy-to-process, smartphones and small robots employ them to increase their sensing capabilities. For example, acceleration sensor data in smartphones is used for healthcare purposes, e.g., recognizing exercise activity in daily situations [1]. Such small robots as Roomba or humanoid type robots use acceleration sensors to identify floor surfaces [2] [3], full-body gesture recognition [4], and person identification [5].

Integration between acceleration and environmental sensors is also a promising approach to increase sensing capabilities and diversity in large environments. For example, acceleration sensors have identified specific person identification among pedestrians using human tracking systems, especially in indoor environments [6] [7].

Using just acceleration sensors remains difficult for accurate positioning due to the accumulated errors, but integrating such environmental systems enables a system to identify and track a person who has an acceleration sensor like a smartphone. Finding one person among many pedestrians will benefit future robot services; for example, a mobile robot can approach a specific user to provide services or change the contents of services based on users.

In this paper, we also propose a method to identify a person who has an acceleration sensor using a human tracking system in real environments. One major unique point of this research is its focus on the changes of accelerations using people's head positions. Past related works, which realized person identification by integrating a human tracking system and acceleration sensors, used people's foot positions to calculate their accelerations to match their acceleration sensor data [6] [7]; such an approach realized accurate acceleration and the walking frequency calculations of people, but its adaption is difficult in such crowded environments as a shopping mall or cluttered environments like a children's playroom. Using the history of head positions to match acceleration sensor data is robust to real situations and the occlusions caused by others or objects in environments.

Our proposed method's evaluation process is one of the strengths of this research. We conducted evaluation experiments in two kinds of real environments with a number of people (Fig. 1). One is a shopping mall, where a relatively large number of people are walking around. The other is a children's playroom, which is cluttered with many objects and small children and a few parents/adults are relatively clumped together. Investigation of the effectiveness of our proposed method through different environments is important to discuss its limitations.

The final publication is available at
http://www.tandfonline.com/doi/full/10.1080/01691864.2015.1095651

Figure 1. Different environments: shopping mall and children's playroom

The rest of our paper is structured as follows. Section II describes related works and clarifies the uniqueness of our work and compares it to them, and Section III introduces our proposed method that uses acceleration sensors to identify people by a human tracking system. Section IV presents our experimental methods, and Section V presents the results. Section VI provides a discussion, and Section VII summarizes our contributions.

2. Related Works

The purpose of this paper is to identify individuals among people in real environments by integrating a human tracking system using environmental depth sensors and acceleration data by wearable devices. To achieve this goal, we need two key techniques: accurate localization of people and robust personal identification in a real environment. Many famous techniques have been proposed for locating people or identifying individuals who possess wearable or mobile personal devices, such as GPS [8], ultrasonic signals [9], UWB [10], and Wi-Fi and/or Bluetooth [11-13]. These devices are prevalent throughout the world, so they can be applied to personal

The final publication is available at
http://www.tandfonline.com/doi/full/10.1080/01691864.2015.1095651

identification in real environments; however, their estimation errors are too large. It is difficult to find an individual who has an acceleration sensor in crowded situations. A method based on processing camera vision also permits positioning with an error of around 5 cm and the identification of individuals [14], but this method is very sensitive to the amount of light. Thus, it is difficult to track and identify people in environments where brightness varies by weather and time, such as a shopping mall. Moreover, privacy issues are raised by ordinary people when a camera network is used in a public space, even though several methods have been proposed which protect privacy in camera sensor networks [15].

To cope with both requirements, i.e., accurate localization and identification, various wearable devices have been developed. The ultra-wide band (UWB) approach is one of the most well-known techniques, since its estimation error is less than 10 cm [10]; however, since it consumes a wide range of radio frequencies, the laws in many countries do not yet allow its use. Another well-known technique is performed with ultrasonic signals in such systems as DOLPHINE [16], Cricket [9], and Active Bat [17]. These devices enable a system to identify people with accurate positioning; however, it is difficult to equip everyday devices with ultrasonic-signal-based positioning since it requires an open space between the receiver and the transmitter.

An alternative approach is to integrate information from multiple sources. Fox et al. proposed a method for tracking multiple people that combines the accuracy benefits of anonymous sensors and ID sensors that consist of infrared and an ultrasound badge [18] [19]. Studies have also combined information from Wi-Fi with other information, such as GSM and GPS [20] [21]. Chen et al. used RFID tag readers with Wi-Fi to improve positioning accuracy and to automatically adapt to changing environmental dynamics [22]. One of the more accurate techniques, developed by Woodman et al., used an accelerometer with Wi-Fi [11] and achieved 50 cm accuracy, 75% of the time. However, these existing techniques have failed to achieve positioning accuracy that can distinguish individuals in environments where many pedestrians are walking.

Similar to our motivation, several researchers have focused on the integration between human tracking systems and wearable sensors to realize accurate localization and identification. Teixeira et al. used inertial sensors on mobile phones to identify and localize multiple people with networked cameras [23]. Sigeta et al. also integrated camera systems and acceleration sensors to identify sensor positions [24]. However, camera systems are sensitive to the amount of light, as described above, and have weaknesses towards occlusions. Past research, which integrated a human tracking system based on laser range finders and acceleration sensors, also have weaknesses towards occlusions because their research works used foot positions for their calculations [6] [7]. We used WiFi-signal strength instead of acceleration sensors and a torso-based human tracking system to be robust for occlusion, but using Wi-Fi signal strength complicates adapting methods to relatively small environments [25].

In this paper we solve these problems using a human tracking system which can estimate people's head positions. To match two kinds of time-series acceleration data from the head position history and wearable devices, we extracted behaviours based on acceleration data and evaluated our proposed system through multiple environments that have different properties.

3. Proposed Method

In this paper, we propose a method that finds a person who is wearing an acceleration sensor in a real environment by integrating the acceleration changes from sensors with the head position information of the people gathered by multiple 3D depth sensors.

3.1 Overview of algorithm

To find sensor holders, our method compares the time-series of the acceleration changes by wearable sensors and the accelerations calculated by trajectories. Fig. 2 shows two people who are moving together in a children's playroom, and the transitions of the accelerations of the X and Z axes were calculated by their trajectories. If multiple people simultaneously move in the same direction, their acceleration values will be similar. As shown in the figure, dealing with Z-axis acceleration increases the possibility of finding differences between a sensor holder and other people. Even if they are not walking around, getting off floors or up from benches will change the accelerations of the Z axis; these changes will be useful to distinguish a sensor holder from the others.

Figure 3 overviews our implemented system. There are two types of sensory information: people's positions tracked by 3D depth sensors and accelerations from wearable sensors. For the predictions, the system calculates the reliabilities by comparing the accelerations calculated by the trajectories and the wearable sensors. It regards the person with maximum reliability as the sensor holder.

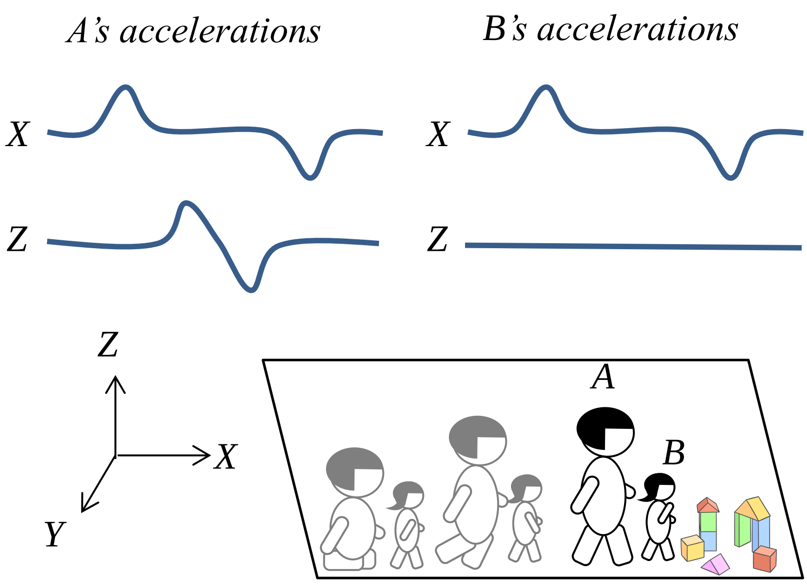


Figure 2. Accelerations of two people calculated by trajectories

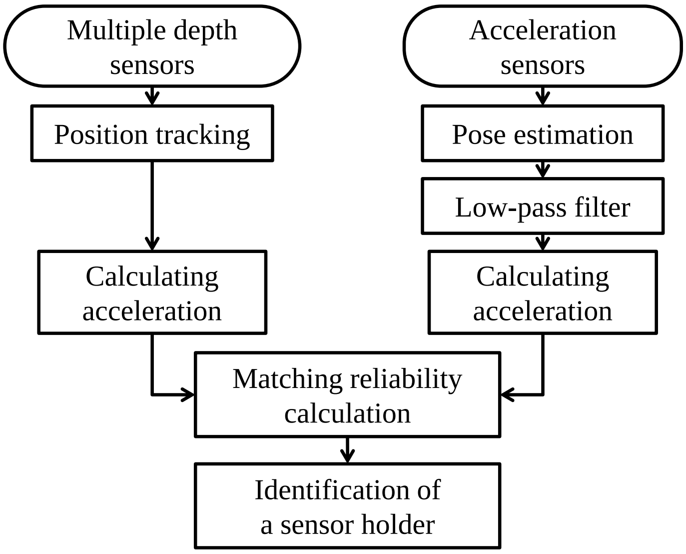


Figure 3. Overview of proposed algorithm

3.2 Acceleration calculated by environmental sensors

To deal with both crowded and cluttered environments, we used the head position histories of people to calculate the acceleration data. For this purpose, we employed environmental depth sensors and an algorithm for human tracking using depth sensors [26] where depth cameras are attached to the ceiling. The tracking system allows us to monitor the position of all the people in the area at 20 Hz with an accuracy of approx. 30 cm. This system has robustness towards illumination changes and the colours of clothing because it uses depth information, which is useful for places where the sun shines in real environments. Since the depth-camera sees from top to down, it is also robust for crowded situations. The time-series of the position data are used to calculate the XYZ accelerations (i.e., $\Delta v/\Delta t$) of the people. In this algorithm, the accelerations, which are calculated by the trajectories of person *i*, are a set of acceleration data:

$$Acc_{pos}(i) = \{acc_{i,0}, acc_{i,1}, \dots acc_{i,n}\} \quad (1),$$

where $acc_{i,0}$ is the acceleration data of X, Y, and Z calculated from the position data, *i* is the anonymous ID assigned by the position tracking system, and *n* is the amount of data from the human tracking system.

3.3 Acceleration calculated by wearable acceleration sensors

We used a TSND121 (ATR-Promotions, 37 (W) x 46 (H) x 12 mm (D)) as a wearable acceleration sensor (Fig. 4, sensor is attached to participant's arm) to measure both the XYZ acceleration and the geomagnetic data at 50 Hz. Note that we used geomagnetic and acceleration data to translate the sensor orientation to the dimensions of the position tracking system and a low-pass filter to decrease the sensor noise. In this algorithm, acceleration data from wearable sensors also comprise a set of acceleration data:

$$Acc_{sensor}(j) = \{acc_{j,0}, acc_{j,1}, \dots acc_{j,m}\}, \quad (2),$$

where $acc_{j,0}$ is the average acceleration data of X, Y, and Z measured by sensor *j* and *m* is the amount of data from the sensor.

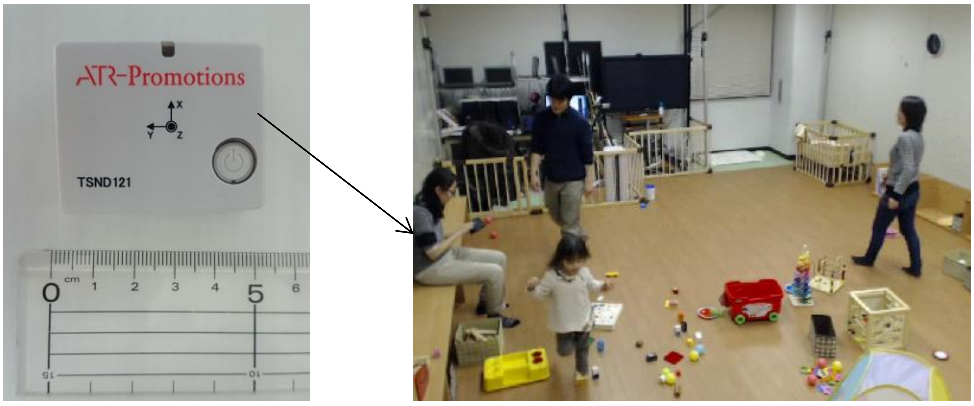


Figure 4. Wearable acceleration sensor (TSND121)

3.4 Reliability calculation

We compared the time-series of the acceleration data from both acceleration and environmental sensors to identify a sensor holder. This comparison can be defined as a measurement of the goodness-of-fit between two sets of time-series data. When the duration of the time-series varies (e.g., in gesture recognition), methods like DP Matching and Hidden Markov Model are most appropriate. On the other hand, our research deals with synchronized time-series data; in such situations other criteria, such as Euclidean distance or Residual Mean Square (RMS), are used to calculate matching reliability. These measurements directly compared both time-series data, and using RMS realizes good estimation for similar purposes [25]. As an another approach, past research, which dealt with synchronized time-series data between laser range finders and acceleration sensors, also directly compared both time-series data by detecting walking behaviours from them [7].

In this paper, we also follow the behaviour-based matching approach, i.e., a way of direct comparison between both time-series data by detecting behaviours in each specific time period. We recognize such rather simple behaviours as moving/stopping in each axis (X, Y and Z) due to the changes of each piece of their data, because our

The final publication is available at
http://www.tandfonline.com/doi/full/10.1080/01691864.2015.1095651

research focuses not only on walking but also on such common behaviours as sitting, standing, and/or playing in real environments. Such irregular movements complicate the analysis of the periodic features of movements, e.g., walking cycles, which were used in past research for person identification. Moreover, the properties of the accelerations calculated by the head positions and the accelerations by wearable sensors are different. For example, accelerations calculated by trajectories are influenced by the noises of the depth information and the delays of the position tracking system, and accelerations calculated by wearable sensors are also influenced by their positions. We believe that dealing with such simple behaviours enables the system to robustly measure the goodness-of-fit between different kinds of time-series data.

First, we estimated the behaviours of each axis from the acceleration data using a threshold at particular time periods:

$$behaviour(acc, axis, t) = \begin{cases} move & th > diff(acc, axis, t, t - t_p) \\ stop & th \leq diff(acc, axis, t, t - t_p) \\ lack & if\ data\ is\ lacked \end{cases} \quad (3)$$

$$diff(acc, axis, t1, t2) = abs(\max(acc(axis, t1, t2)) - \min(acc(axis, t1, t2)), \quad (4)$$

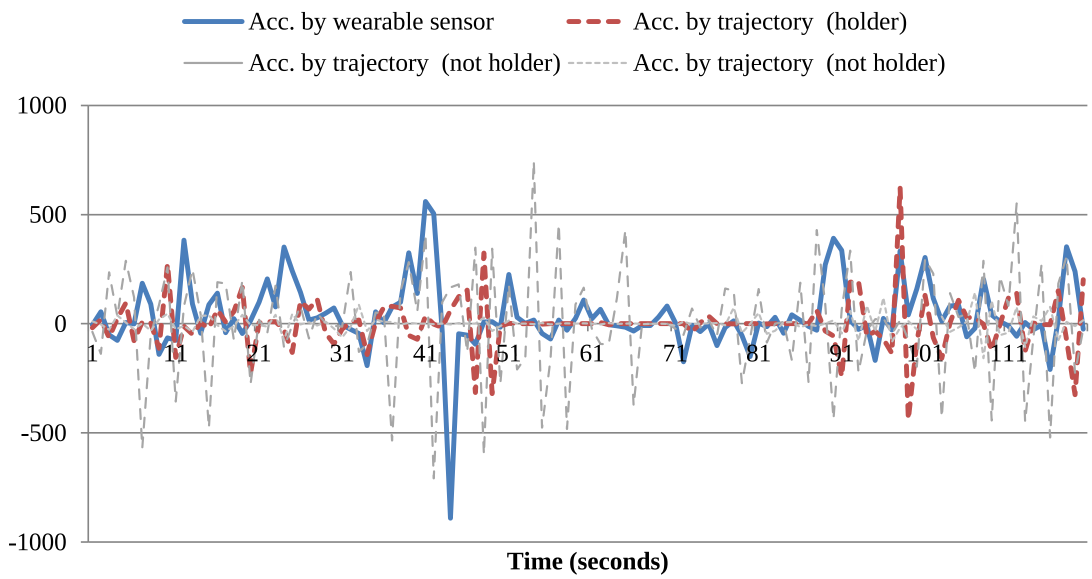
where *th* is the threshold value, t_p is the time period to extract the acceleration data, *axis* identifies which axis will be used for calculation, and *max* and *min* functions return maximum and minimum values within the acceleration data between *t1* and *t2*. Fig. 5 shows an example of the Z-axis acceleration raw data and processed data through Eq. (3). These values for all axes are compared to calculate the ratio of the matched data as the sensor holder's reliability:

$$reliability(i, j) = \frac{\sum_{t=tp}^{n} \sum_{axis=x,y,z} match(\ behaviour(Acc_{pos}(i), axis, t), behaviour(Acc_{sensor}(j), axis, t)}{Number\ of\ axes} \quad (5),$$

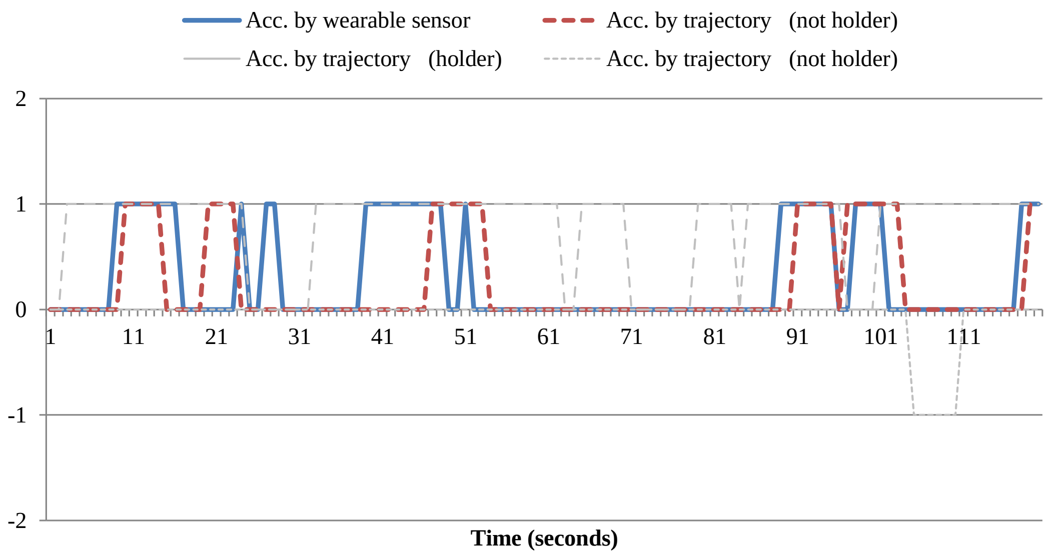
$$match\ (behaviour\ A, behaviour\ B) = \begin{cases} \propto & if\ \ behaviour\ A == behaviour\ B \\ 0 & if\ \ behaviour\ A\ != behaviour\ B \\ -\propto & if\ behaviourA\ or\ behaviourB = lack \end{cases} \quad (6).$$

The final publication is available at
http://www.tandfonline.com/doi/full/10.1080/01691864.2015.1095651

We designed a *match* function to return a negative coefficient when either datum is missing at the time period to decrease a candidate's reliability. This is needed to exclude a candidate with small number and well-matched data towards the sensor holder's data. The reliability value is divided by the number of axes to normalize the reliabilities; in this research we use three bits of acceleration data: X, Y, and Z. The system calculates the reliability for all the combinations among all of the wearable sensors and the tracked people and extracts the person with maximum reliability as the sensor holder. Note that the parameters of each method (i.e., th, t_p, and α) will be tuned for each environment by a grid search to realize the best performance.



(a) Z-axis acceleration (raw)



(b) Z-axis acceleration (results from *behaviour* function, 1: move, 0: stop, -1: missing)

Figure 5. Wearable acceleration sensor (TSND121)

4. Data Collection

In this research, we gathered both acceleration and environmental sensor data at two real environments that have different properties. One is a shopping mall where many people are walking in a relatively large environment. The other is a children's playroom where a few families are playing together in a relatively small environment. We believe that investigating the performances of our proposed method in such different environments is critical for understanding its limitations.

4.1 Environment 1: shopping mall

Figure 6 shows a shopping mall where we installed 49 3D range sensors in the ceiling (combining a Panasonic D-Imager, an ASUS Xtion, and a Velodyne HDL-32E) in an area that covered around 900 m^2. This mall's arcade includes restaurants, cafes, and miscellaneous stores. People pass directly through the arcade, and others visit shops. During lunch time, people frequently rested on the benches in the central square.

The final publication is available at
http://www.tandfonline.com/doi/full/10.1080/01691864.2015.1095651

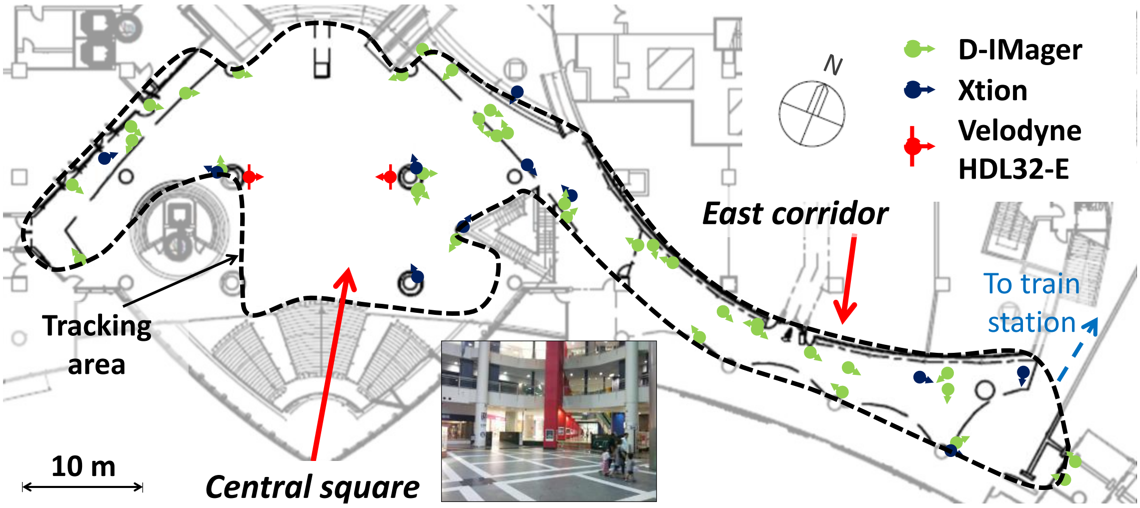


Figure 6. Shopping mall environment

4.2 Environment 2: children's playroom

Figure 7 shows a children's playroom where we installed 3D depth sensors. We used an 11 ASUS Xtion Live pro for position tracking in this environment, and the total covered area was around 40 m^2. Toys, books, and chairs are available for visitors who can freely play in the room, including parents with their children, but they did not move around much. They often sat for relatively long periods and watched their children.

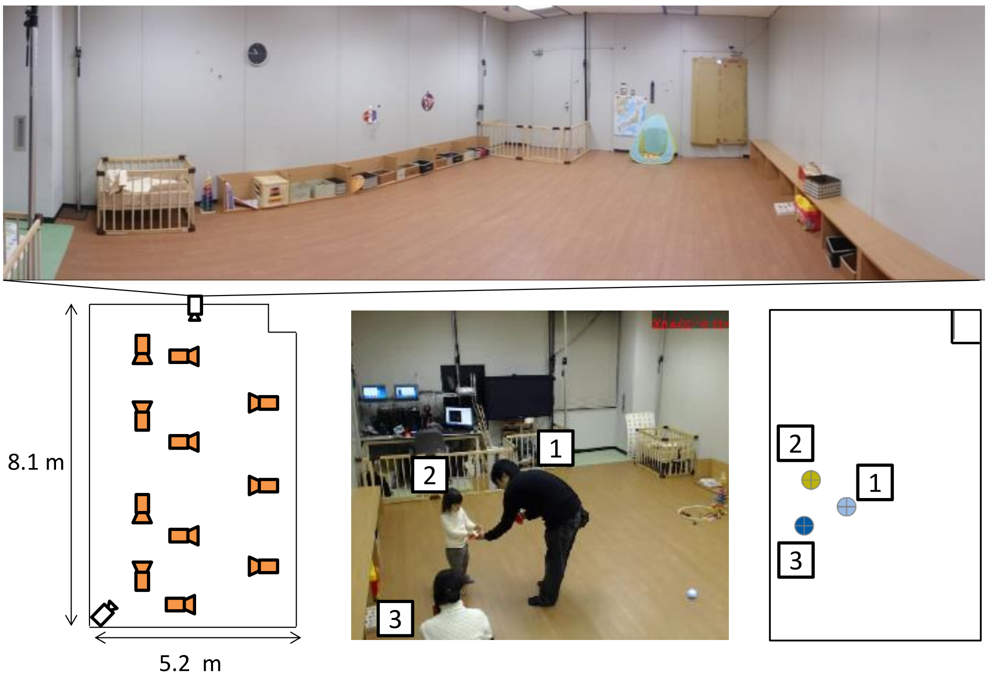


Figure 7. Children's play environment

The final publication is available at
http://www.tandfonline.com/doi/full/10.1080/01691864.2015.1095651

4.3 Procedure

We chose different procedures for each environment because of their different properties. In the shopping mall, an experimenter wore two acceleration sensors on his dominant arm and his hip pocket, and walked from the entrance to the square (or reverse) five times and strolled through the square five times at lunch time to gather sensor data. The purpose of using two acceleration sensors is to investigate the tendency of the proposed method toward a use of multiple sensors simultaneously, and the different position of the sensor. In total, we conducted ten trials (each was two minutes) for the data collection in this environment, which is usually used by ordinary people like office workers, families, and couples.

Different from the shopping mall experiment, in the children's playroom, the parents wore acceleration sensors on their dominant arms and behaved as if they were visiting a standard children's playroom; the experimenter did not wear an acceleration sensor. We also conducted ten trials (each was two minutes) for the data collection in this environment. Basically, one or two families remained in it during the data collection.

5. Evaluation

5.1 Performance evaluation with alternative methods

In this section, we evaluate our proposed method by comparing two alternative methods: Euclidean distance with a behavioural method and the earth mover's distance (EMD)-based method. The evaluations were conducted with offline.

The final publication is available at
http://www.tandfonline.com/doi/full/10.1080/01691864.2015.1095651

The Euclidean distance method is commonly used for comparing two kinds of time-series data sets. This is relatively simple measurement but it is often used as a basic criterion to investigate the performances of new methods. Therefore, we also employed this measurement for comparison. Here, we defined a behavioural three-dimensional vector using the behaviours of each axis at each time step, like *state* (*i*, *t*) = (*move(1)*, *move(1)*, *stop(0)*), to calculate the Euclidean distances instead of raw acceleration data. Therefore, the characteristics of this alternative method resemble the proposed method, and the main difference is the reliability function. This comparison is important to investigate the effectiveness of our reliability function to identify sensor holders.

The EMD-based method uses EMD, which measures the distance between two probability distributions [27], as a matching reliability between two time-series data, instead of the reliability calculated by Eq. (5). This metric is often used to compute the distance between such probability distributions as colour histograms. We think it is similar to the comparison compare between acceleration data as shown in the figure 5, therefore we employed this measurement too. Similar to the Euclidean distance with a behavioural method, we used behavioural three-dimensional vectors. Therefore, the characteristics of this alternative method deal with probabilistic data without time considerations, unlike our proposed method which considered time-synchronization. We calculated the probability distributions based on the frequency for all the time-series data and calculated EMD to find the sensor holder. The system extracts the person with a minimum EMD as the sensor holder.

Figure 8 shows the performance of the proposed method by using the dominant arms' acceleration sensors as well as the alternative methods. The parameters of each method (i.e., th, t_p, and α) are tuned for each environment by a grid search to realize the best performance. Our proposed method achieved 85% person identification and outperformed the alternative methods (Euclidean distance with behavioural method: 70%, EMD-based method: 65%).

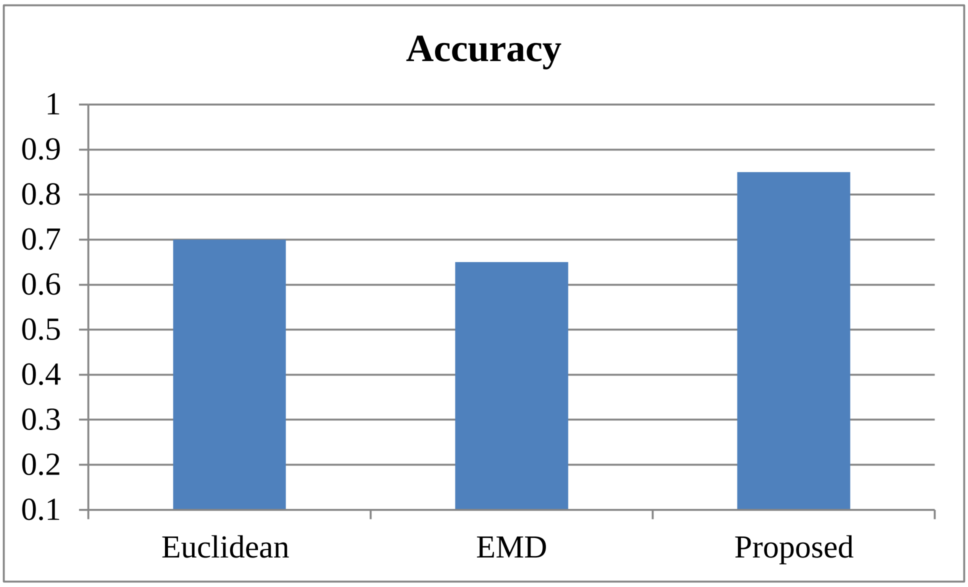


Figure 8. Averaged performance in both environments

We also show the performances in each environment. The performance in the mall environment is shown in Fig. 9. On average, 286.3 people, including an experimenter, were detected by the human tracking system through data collection in the mall (chance rate was 100/286.3=0.35%). The proposed method achieved 80% person identification; its performance was also better than the alternative methods (Euclidean distance with behavioural method: 60%, EMD-based method: 40%). Fig. 10 shows the X-Y trajectories gathered from the shopping mall environment during a trial. In this trail, there are 295 people are detected; the proposed method and the Euclidean method succeeded to identify the experimenter among them, but the EMD method failed to identify. The red (or black) bold trajectory showed the sensor holder. The blue (or black) trajectory showed the next candidate by proposed method, and the first candidate by EMD method. The reason of similar reliabilities is because they moved to left side from right side at similar timing at the first half time period, even though the last time period's trajectories are different. The sensor holder's value of Eq. (5) in the proposed method was 0.66, and the next candidate's value of Eq. (5) in the proposed method was 0.51. On the other hand, the sensor holder's EMD value was 0.30, and the next candidate's value was 0.24. In this case, both people stay around during few seconds in the middle of data collection; even if the timing of these stop behaviours are different, EMD method did not consider time-synchronization then the method misidentified the sensor holder.

Fig. 11 shows the case where the proposed method failed to identify the sensor holder (red). In this case, the EMD method only could identify the sensor holder (blue). This candidate is the next candidate of the proposed method. There trajectories and time periods are similar, only the proposed method failed to identify due to noise of detected positions or delay. The sensor holder's reliability by the proposed method become lower because behaviours were not well fit between sensors within certain time periods; but the total number of behaviours were similar then EMD method succeed to identify the sensor holder.

However, in total, even though there are several similar trajectories towards the sensor holder, the system successfully found him through the proposed method.

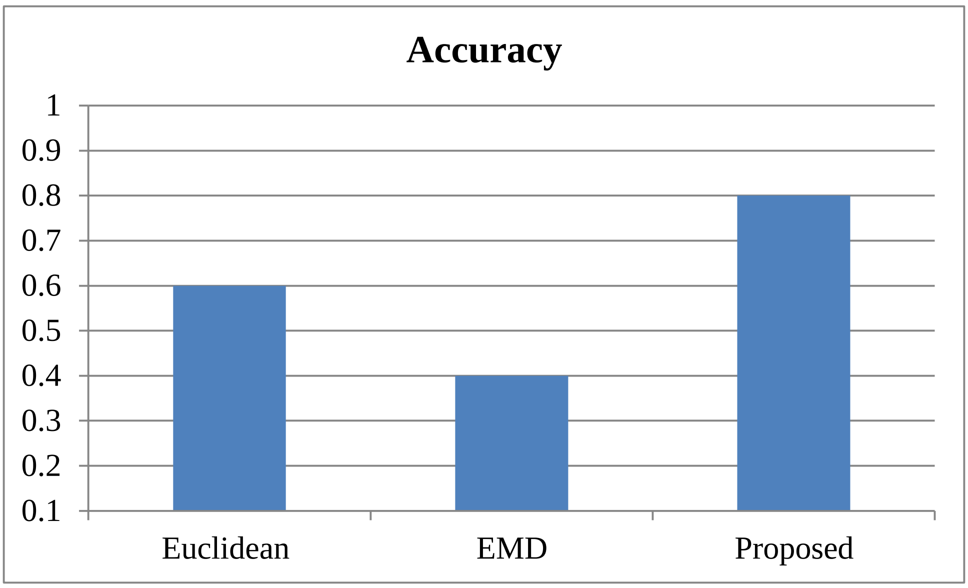


Figure 9. Performance in shopping mall environment

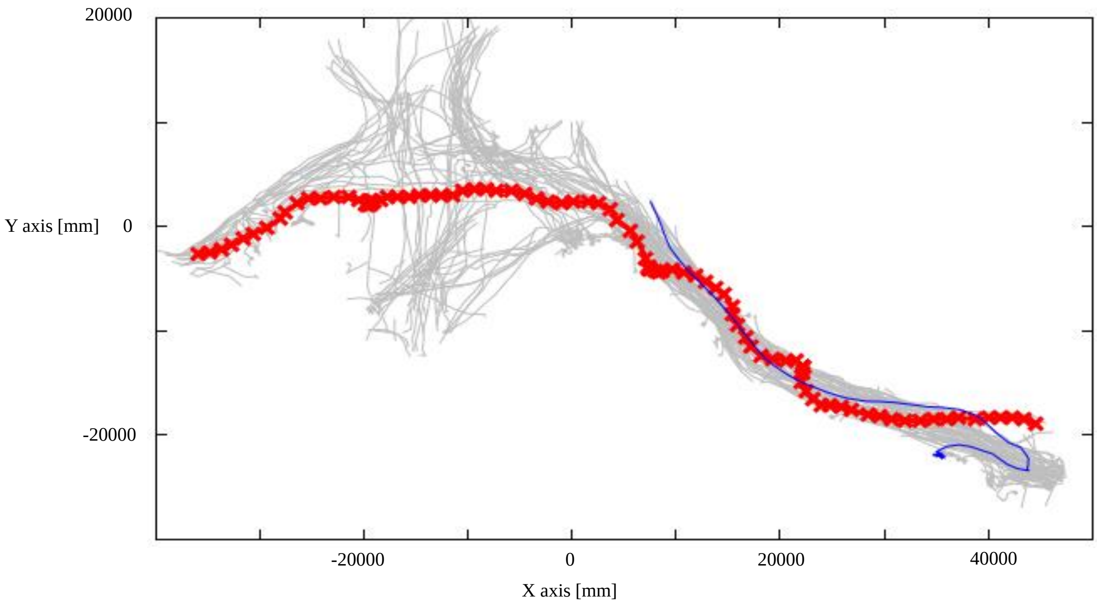


Figure 10. Trajectories at a shopping mall: sensor holder’s and second candidate trajectories are emphasized.

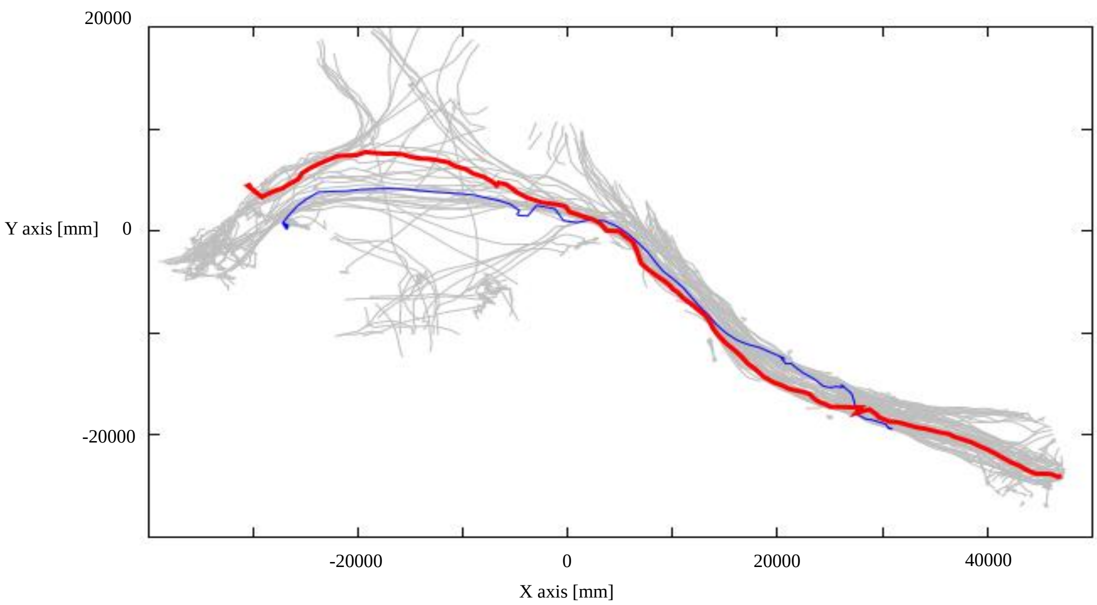


Figure 11. A case where the proposed method failed to identify the sensor holder

The performance in the children's playroom environment is shown in Fig. 12. On average, 4.8 people were simultaneously detected by the human tracking system through data collection in it (chance rate was 100/4.8=20.8%). The proposed and EMD-based methods achieved 90% person identification; the Euclidean distance with a behavioural method achieved 80% person identification. Fig. 13 shows the X-Y trajectories gathered from the children's playroom environment. The system identified the sensor holder who did not move around very much (red or black bold trajectory), even if other people did not move around like the sensor holder (green or gray bold trajectory). Even though they are similar, all of method could identify the sensor holder. The sensor holder's value of Eq. (5) in the proposed method was 0.90, and the next candidate's value of Eq. (5) in the proposed method was 0.81. Moreover, the sensor holder's EMD value was 0.13, and the next candidate's value was 0.24. The main reason of why all methods could identify the sensor holder is a use of Z-axis information; the timing of standing behaviors are different between them, even if their X-Y trajectories are quite similar.

Since the number of trajectories is smaller than the shopping mall environment, their differences are clear. We assumed that a smaller environment would be more difficult for person identification because their movements would be limited, but the performances of the alternative methods were similar to the proposed method.

The evaluation results show that our proposed method achieved more accurate identification than the alternative methods when the system needs to deal with various environments and demonstrated the concept's potential through evaluations in real environments.

The final publication is available at
http://www.tandfonline.com/doi/full/10.1080/01691864.2015.1095651

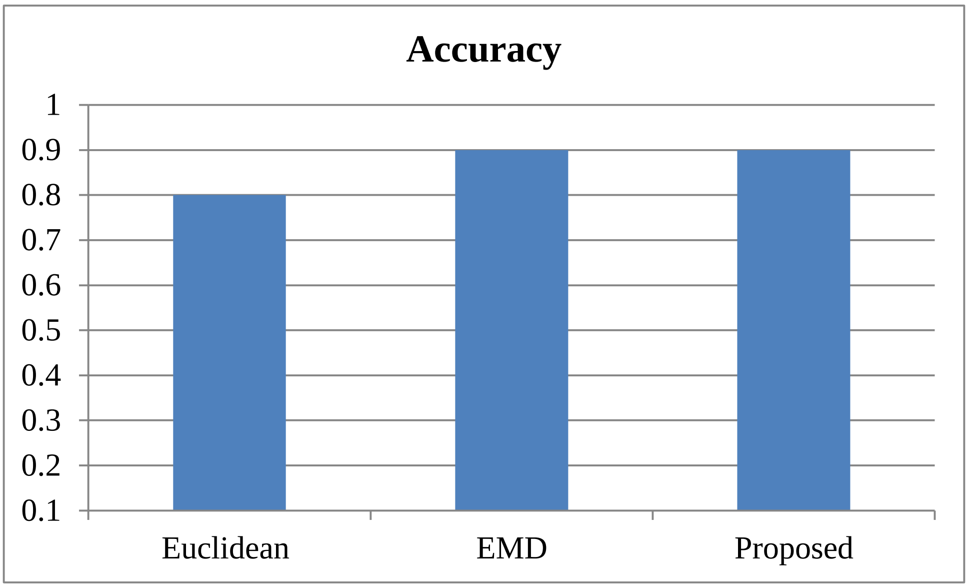


Figure 12. Performance in children's playroom environment

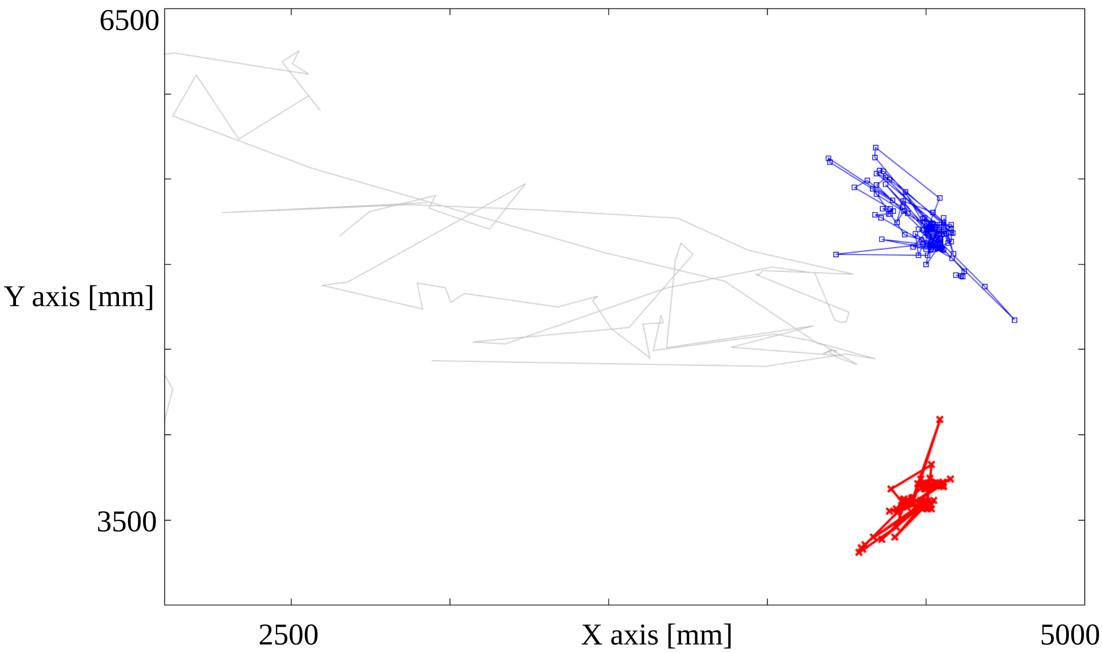


Figure 13. Trajectories in children's playroom environment. Red (right-bottom, bold-cross) indicates trajectory of sensor holder, and blue (right-top, square) indicates a person with a similar trajectory as sensor holder.

Moreover, we investigate the performance of the proposed method by using the hip pocket's acceleration sensors. As a result, our proposed method achieved 90% person identification, which outperforms the other methods using acceleration sensors on dominant arms. The reason of improved performance of identification would be caused by less time gap of acceleration peaks between head positions and sensor data. When the sensor holder placed the sensor to her/his dominant arm and swung the arm largely, the peak timing of acceleration data will be different from the head acceleration data. If the sensor is attached to a wristwatch, the performance also becomes lower due to the effects of time delay. On the other hand, if the sensor holder placed the sensor to the hip pocket, such effects would be small then the identification performance would increase. In total, the proposed method showed higher performances for person identification compared with alternative methods. The additional evaluation showed the tendency of the proposed method towards the sensor position, and robustness towards a use of multiple sensors simultaneously. Note that this situation is a little bit different from a situation where multiple sensor holders are exist because the same person has two sensors, but these results would show that the proposed method could deal with multiple sensors.

5.2 Analysis of failure cases in proposed method

Here we investigate why our proposed method failed to identify three of the 20 persons with wearable acceleration sensors: two in the mall and one in the children's playroom. In the mall environment, there are two reasons: swapping of the ID of the human tracking system because of the cluttered situations and the ID's disappearance for a few seconds. When the experimenter crossed a group of people at a narrow place, the human tracking system wrongly assigned the ID and could not calculate the correct behaviours of the experimenter, and therefore the identification failed. In another failure case, the tracking system lost the experimenter for a few seconds; even if it correctly re-assigned the same ID after the person's disappearance, it decreased the reliability of the person's identification. In fact, the trajectory of the person whom the system misidentified resembled the correct sensor holder.

In the children's playroom environment, insufficient coverage of the depth sensors caused incorrect identification. Because of a lack of sensor cover area, the person's height was wrongly estimated. Since the system incorrectly assigned behaviours even when the person's height did not change, this error reduced the calculation's reliability.

In both environments the main reason for the failures was the position tracking system and its settings. But this problem might be solved using more accurate sensor systems because occlusions and disappearances were mainly caused by noise or the low density of the sensors.

6. Discussion

6.1 Possible applications

A system, which features accurate positioning and person identification, presents a wide range of useful applications, including such services as welcoming specific customers or guidance around facilities using a mobile robot. In fact, our past work demonstrated a concept of such services with a mobile robot in a real environment [25].

In addition to that developed application, another possibility is applying our technique to ambient intelligent environments in which facilities (displays, music, illumination, etc.) are actively controlled based on pre-registered personal information. For instance, an electronic poster could adjust its advertising content in advance to target material to a particular person.

Another major future application is as a tool in marketing services. Shops are looking for people's behavioural data in relation to purchases. In return for access to a user's personal information (purchase and location history), they could offer the free use of such physical services as self-navigating shopping carts and drink delivery robots.

6.2 Identification of multiple sensor holders

In this paper, we considered a situation where only one sensor holder exists in the environment. We believe that our proposed method can be extended to simultaneously find multiple sensor holders, because it was designed to calculate the matching reliability between each acceleration sensor and each person. In real-time applications, sensor data will be sent by Wi-Fi, and sensor data delays might increase error calculation, but the number of sensors is not related to critical errors in our proposed method. One concern is the overlapping of candidate results between different sensor holder estimations; in such situations, the history of matching reliability is useful to determine who is a sensor holder. Although such a mechanism hasn't been implemented in our proposed method yet, future work will address it.

6.3 A use of acceleration data from other body part

In this paper, we compared acceleration data of different body part e.g., head and dominant arm/hip pocket. If the system compared the acceleration data of the same parts, the performance would be improved. However, there are two kinds of difficulties to compare the acceleration data of the same parts: robust tracking of body parts in large/crowded situations and estimating body parts where the sensor is attached. In fact, large and/or crowded situations, robust tracking of body parts were difficult with our system, even if it could track people. Past research work achieved the estimation of acceleration sensor position in camera image, but for this purpose the sensors need to cover all body parts of all people in an environment. In real settings, it would be difficult to completely cover all body parts of all people by sensors. Thus, the main reason of why we used head acceleration is because this part is relatively easy to detect than other parts in a real situation. Therefore, in such situation, we can say that a use of head acceleration would be effective than other body parts from the viewpoints of continuous tracking; note that we did not claim the advantage of comparing acceleration data between different body parts.

From these reasons, we tried to use a kind of abstracted acceleration data, as defined by Eq. (3), to identify people by using head acceleration data and other body part's acceleration data. An idea of using abstracted sensor data might be popular for acceleration data, e.g., recent pedometers can robustly count user's step in spite of its positions by using (abstracted) acceleration only. In other words, in real settings, we used abstracted acceleration data from different body parts, as one of realistic solutions to achieve to integrate a human tracking system and acceleration sensors data.

6.4 Privacy concerns

Our system provides accurate positioning with acceleration sensors, which is a useful technique if used appropriately. However, there has been discussion about privacy concerns with positioning systems (e.g., [28]), and unfortunately, our method increases privacy risks.

One question is whether positioning should be conducted at a central server or at each client device (a decentralized approach). The latter has an advantage for privacy (e.g., [29, 30]). Our approach is the former, which increases privacy risks. However, even if positioning is conducted on client devices such as a smartphone, the user's position data must still be sent to the system for such physical services as the applications discussed in this research. In such situations, a user must clearly understand that the system knows her position because a physical interaction happened at her position based on her request. Here, another important design consideration is to obtain consent from the users to exploit such information.

6.5 Limitations

We acknowledge other limitations. Our proposed method provided 85% accurate identification in real environments, but we did not compare its performances with such state-of-the-art time-series comparison algorithms as DUST [31] and SAX [32]. Using these methods might increase performance and decrease the required data length. Since this algorithm identifies a person by focusing on head positions, not only XY trajectories, our method failed to distinguish people with quite similar movements. Long-time observation might enable the system to identify them, but that increases the data length. However, we believe that this technique can provide various physical services in daily life for registered people, particularly in public commercial spaces that are concerned about privacy.

We investigated the performances with multiple sensors, but it is a not situation where multiple sensor holders exist. One of major problem with multiple sensor holders is that the reliabilities of different sensors were the same person. To solve such problem, a simple approach is that the system simply assign a sensor which has highest reliability to the person to solve the problem.

7. Conclusion

This paper proposed a person identification method for those wearing an acceleration sensor in real environments using the histories of head positions estimated by environmental depth sensors. The system compared two kinds of time-series acceleration data: those calculated by trajectories and those by wearable sensors. For comparison, we proposed a behaviour-based matching algorithm to compare the timing of the changes of each bit of acceleration data at certain time periods instead of directly comparing the acceleration values.

We investigated the performance of our proposed method through data collection in two kinds of real environments: a shopping mall and a children's playroom. We gathered 40 minutes of sensor data with 20 trials, and our proposed method correctly realized 85% identification. It also showed better results than alternative methods that used other metrics. We also investigated the tendency of our proposed method through additional data collection where the experimenter placed acceleration sensors in different places.

Acknowledgements

This research was supported by the Strategic Information and Communications R&D Promotion Programme (SCOPE), Ministry of Internal Affairs and Communications (132107010). We thank the administrative staff at the Asia and Pacific Trade Center for their corporation. We also thank the children, parents, and playroom staff for their helpful participation.

References:

[1] J. Yang, "Toward physical activity diary: motion recognition using simple acceleration features with mobile phones," in Proceedings of the 1st International Workshop on Interactive Multimedia for Consumer Electronics, pp. 1-10, 2009.

[2] P. Giguere, 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.

[3] R. Matsumura, M. Shiomi, T. Miyashita, H. Ishiguro, and N. Hagita, “What kind of floor am I standing on? Floor surface identification by a small humanoid robot through full-body motions,” *Advanced Robotics*, 2015 (to appear)

[4] M. Cooney, T. Kanda, A. Alissandrakis, and H. Ishiguro, “Designing Enjoyable Motion-Based Play Interactions with a Small Humanoid Robot,” *International Journal of Social Robotics*, vol. 6, no. 2, pp. 173-193, 2014.

[5] R. Matsumura, M. Shiomi, T. Miyashita, H. Ishiguro, and N. Hagita, “Who is Interacting With me?; Identification of an Interacting Person Through Playful Interaction With a Small Robot,” *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 2, pp. 169-179, 2014.

[6] T. Ikeda, H. Ishiguro, T. Miyashita, and N. Hagita, “Pedestrian identification by associating wearable and environmental sensors based on phase dependent correlation of human walking,” *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-10, 2013.

[7] K. Morioka, F. Hashikawa, and T. Takigawa, “Human Identification Based on Walking Detection with Acceleration Sensor and Networked Laser Range Sensors in Intelligent Space,” *International Journal on Smart Sensing and Intelligent Systems*, vol. 6, no. 5, pp. 2040-2054, 2013.

[8] E. D. Kaplan, and C. J. Hegarty, *Understanding GPS: Principles and Applications*: Artech House, 2006.

[9] N. B. Priyantha, A. K. L. Miu, H. Balakrishnan, and S. Teller, “The cricket compass for context-aware mobile applications,” in Proceedings of the 7th annual international conference on Mobile computing and networking, Rome, Italy, pp. 1-14, 2001.

[10] C. Woo Cheol, and H. Dong-Sam, “An accurate ultra wideband (UWB) ranging for precision asset location,” in 2003 IEEE Conference on Ultra Wideband Systems and Technologies, pp. 389-393, 2003.

[11] O. Woodman, and R. Harle, “Pedestrian localisation for indoor environments,” in Proceedings of the 10th International Conference on Ubiquitous Computing, Seoul, Korea, pp. 114-123, 2008.

[12] A. Baniukevic, C. S. Jensen, and L. Hua, “Hybrid Indoor Positioning with Wi-Fi and Bluetooth: Architecture and Performance,” in 2013 IEEE 14th International Conference on Mobile Data Management (MDM), pp. 207-216, 2013.

[13] W. Yapeng, Y. Xu, Z. Yutian, L. Yue, and L. Cuthbert, “Bluetooth positioning using RSSI and triangulation methods,” in 2013 IEEE Conference on Consumer Communications and Networking Conference (CCNC), pp. 837-842, 2013.

[14] S. Lao, and M. Kawade, "Vision-Based Face Understanding Technologies and Their Applications," *Advances in Biometric Person Authentication*, Lecture Notes in Computer Science S. Li, J. Lai, T. Tan *et al.*, eds., pp. 339-348: Springer Berlin Heidelberg, 2005.

[15] T. Winkler, and B. Rinner, “Security and Privacy Protection in Visual Sensor Networks: A Survey,” *ACM Comput. Surv.*, vol. 47, no. 1, pp. 1-42, 2014.

[16] M. Minami, Y. Fukuju, K. Hirasawa, S. Yokoyama, M. Mizumachi, H. Morikawa, and T. Aoyama, "DOLPHIN: A Practical Approach for Implementing a Fully Distributed Indoor Ultrasonic Positioning System," *UbiComp 2004: Ubiquitous Computing*, Lecture Notes in Computer Science N. Davies, E. Mynatt and I. Siio, eds., pp. 347-365: Springer Berlin Heidelberg, 2004.

[17] A. Ward, A. Jones, and A. Hopper, “A new location technique for the active office,” *IEEE Personal Communications*, vol. 4, no. 5, pp. 42-47, 1997.

[18] D. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello, "Bayesian filtering for location estimation," *IEEE pervasive computing*, vol. 2, no. 3, pp. 24-33, 2003.
[19] D. Schulz, D. Fox, and J. Hightower, "People tracking with anonymous and ID-sensors using Rao-Blackwellised particle filters," in Proceedings of the 18th International Joint Conference on Artificial Intelligence, Acapulco, Mexico, pp. 921-926, 2003.
[20] A. LaMarca, J. Hightower, I. Smith, and S. Consolvo, "Self-mapping in 802.11 location systems," *UbiComp 2005: Ubiquitous Computing*, pp. 87-104: Springer, 2005.
[21] J. Hightower, S. Consolvo, A. LaMarca, I. Smith, and J. Hughes, "Learning and recognizing the places we go," in Proceedings of the 7th International Conference on Ubiquitous Computing, Tokyo, Japan, pp. 159-176, 2005.
[22] Y.-C. Chen, J.-R. Chiang, H.-h. Chu, P. Huang, and A. W. Tsui, "Sensor-assisted wi-fi indoor location system for adapting to environmental dynamics," in Proceedings of the 8th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems, pp. 118-125, 2005.
[23] T. Teixeira, D. Jung, and A. Savvides, "Tasking networked CCTV cameras and mobile phones to identify and localize multiple people," in Proceedings of the 12th ACM International Conference on Ubiquitous Computing, Copenhagen, Denmark, pp. 213-222, 2010.
[24] O. Shigeta, S. Kagami, and K. Hashimoto, "Identifying a moving object with an accelerometer in a camera view," in IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008, pp. 3872-3877, 2008.
[25] M. Shiomi, K. Kurumizawa, T. Kanda, H. Ishiguro, and N. Hagita, "Finding a person with a Wi-Fi device in a crowd of pedestrians," *Advanced Robotics*, vol. 28, no. 7, pp. 441-448, 2014.
[26] D. Brscic, T. Kanda, T. Ikeda, and T. Miyashita, "Person Tracking in Large Public Spaces Using 3-D Range Sensors," *IEEE Transactions on Human-Machine Systems*, vol. 43, no. 6, pp. 522-534, 2013.
[27] O. Pele, and M. Werman, "Fast and robust earth mover's distances," in 2009 IEEE 12th international conference on Computer vision, pp. 460-467, 2009.
[28] L. Huang, H. Yamane, K. Matsuura, and K. Sezaki, "Towards modeling wireless location privacy," in Proceedings of the 5th International Conference on Privacy Enhancing Technologies, Cavtat, Croatia, pp. 59-77, 2006.
[29] P. Bellavista, A. Corradi, and C. Giannelli, "Efficiently managing location information with privacy requirements in Wi-Fi networks: a middleware approach," in 2nd International Symposium on Wireless Communication Systems, pp. 91-95, 2005.
[30] J. I. Hong, G. Boriello, J. A. Landay, D. W. McDonald, B. N. Schilit, and J. D. Tygar, "Privacy and Security in the Location-enhanced World Wide Web " in Proceedings of Fifth International Conference on Ubiquitous Computing: Ubicomp, pp., 2003.
[31] S. R. Sarangi, and K. Murthy, "DUST: a generalized notion of similarity between uncertain time series," in Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, pp. 383-392, 2010.
[32] A. Camerra, T. Palpanas, J. Shieh, and E. Keogh, "iSAX 2.0: Indexing and mining one billion time series," in 2010 IEEE 10th International Conference on Data Mining (ICDM), pp. 58-67, 2010.