Detecting Room-to-Room Movement by Passive Infrared Sensors in Home Environments

Kazuya Murao[1], Tsutomu Terada[1], Ai Yano[2], and Ryuichi Matsukura[3]

[1] Graduate School of Engineering, Kobe University
1-1 Rokkodai, Nada, Kobe, Hyogo 657-8501, Japan
{murao,tsutomu}@eedept.kobe-u.ac.jp
[2] Fujitsu Laboratories Ltd.
4-1-1 Kamikodanaka, Nakahara, Kawasaki, Kanagawa 211-8588, Japan
yano.ai@jp.fujitsu.com
[3] Fujitsu Limited
4-1-1 Kamikodanaka, Nakahara, Kawasaki, Kanagawa 211-8588, Japan
r.matsukura@jp.fujitsu.com

Abstract. We discuss in this paper the problem of simultaneous tracking, which exploits the synergy between location and movement to provide the information necessary for intelligent home appliance control. Our goal is to carry out accurate movement estimation for multiple people in a home environment. We propose an approach that uses information gathered using only passive infrared sensors commonly found in lighting control systems. No special devices or video cameras are needed. Moreover, it is not necessary to carry out data collection for training. We evaluated our approach by conducting experiments in a real home fitted with sensors and we confirmed that room-to-room movement was detected with an accuracy of 0.82 for two inhabitants who moved freely through the house.

1 Introduction

The Kyoto Protocol was initially adopted in Kyoto in 1997 and it was entered into force in 2005. Under the protocol, 37 countries commit themselves to a reduction of greenhouse gases, and all the member countries give general commitments. In the negotiations, 37 countries including the US collectively agreed to reduce their greenhouse gas emissions by 5.2% on average for the period 2008-2012. Reduction of green effect gas emissions has been recognized as one of the social goals of the world. Moreover, the 2011 earthquake off the Pacific coast of Tohoku, also known as the 2011 Tohoku earthquake, or the Great East Japan Earthquake, occurred on 11 March 2011. The earthquake triggered powerful tsunami waves that reached heights of up to 40.5 m (133 ft). The tsunami struck and severely damaged nuclear power plants, which caused a number of nuclear accidents. This shattered the myth that nuclear power is a safe energy resource. As of April 2012, 53 of the 54 nuclear power plants had been disabled or taken offline, so Japan is facing some serious power shortages[4].

[4] http://www.gengikyo.jp/english/status/ChartOfPowerPlant.html

Home Energy Management System (HEMS) and Building Energy Management System (BEMS) have attracted a great deal of attention due to the growing interest in energy conservation around the world and the development of related technologies. These systems manage home appliances and optimize power consumption by using sensor hardware. For example, an infrared sensor in a TV or lighting system detects the absence of humans and reduces power consumption by turning these devices off when no one is there. However, simple management such as "turn on when someone comes" and "turn off when no one is there" may not be appropriate for other types of appliances, such as for air-conditioners and broadcast satellite tuners. For example, if an air-conditioner is simply turned on and off, the aironditioning will stop even when a user temporarily leaves the room to go to the lavatory. This will lead to an increase in temperature, and the user will have to turn on the air-conditioner again, which will consume excess power. A TV can be turned off, but a broadcast satellite tuner cannot be turned off as easily since it takes longer to activate itself. Appliances that take longer to activate should be managed before the user arrives. Smart management without taking human movement into consideration just deteriorates quality of life (QoL). Intelligent management of appliances requires information on human movement; this has been acquired through systems using cameras or radio frequency identification (RFID) tags in conventional studies. Such systems, however, do not consider privacy issues, and they involve additional tasks for the users such as carrying around a device all day long. A system for general use should not require users to carry around or operate any devices, and the systems should not cause stress to the users when monitoring them.

We propose a method for detecting room-to-room movements of inhabitants in a home environment by using infrared sensors. Our method is appealing because of the following points;

- Users do not have to carry or operate any devices.
- The number of infrared sensors needed is relatively small ($\approx 5\ m^2$/sensor).
- It is not necessary to collect training data.

We adopted the infrared sensor since it has certain advantages, as follows:

- Infrared sensors are relatively cheap ($\approx$ 3 USD).
- Infrared sensors are often already installed in homes to control lighting.
- Infrared sensors do not invade our privacy since only in/out information is captured

Our goal in this study is to detect room-to-room movements of multiple inhabitants even when they moved at the same time. In addition, we investigated the importance of sensor position by changing sensor combinations.

The paper is organized as follows. Section 2 introduces related work. Section 3 presents the assumed environment. We describe the proposed method in Section 4 and discuss its performance in Section 5. Finally, Section 6 concludes the paper.

2 Related Work

Much research on human movement detection using sensors has been done. Many methods use cameras, which can track humans and detect invaders, because they obtain a great deal of information through images, video, and sound. Two of the main purposes of systems using cameras are to detect invaders from outside and to monitor workers in factories. Therefore, these systems create a feeling of being kept under surveillance for inhabitants, which is not appropriate for home use from the viewpoint of privacy[1].

A human tracking system using RFID has been proposed[2]. Users with an RFID reader are traced by touching RFID tags attached to objects as they move. The advantage of the system is that transitions are correctly detected unless the user forgets to touch the RFID reader. Although the system is easily applicable to office environments where ID cards with RFID tags are used as keys to enter rooms, it is a difficult constraint for home users to have to carry a card and touch a reader whenever they move around.

A position estimation method using dead reckoning with inertia sensors has also been proposed[3]. The system is complementarily used with a global positioning system (GPS) to obtain location information when the GPS does not work well in the shade of buildings and trees or when underground. The system is assumed to be used in a broad area, as errors accumulate as time elapses in narrow areas where the user performs fragmentary movements and pivot turns (changing direction on the spot). Moreover, users have to attach sensors on the fixed position whenever they move, which is unrealistic and unacceptable in a home environment.

Position and transition detection systems using ubiquitous sensors without any portable devices have also been proposed. One example is a door-level movement detection method using pressure sensors attached to central heating, ventilation, and air conditioning (HVAC) management systems[4]. Sensors attached to an air filter detect room transitions throughout the house from the changes in air pressure that occur when people pass through doorways and open and close doors. This kind of system is easy to install and maintain since sensors are attached to only one filter. However, the detection accuracy is not more than 65%, and much depends on the floor plan. As of 1997, approximately 66% of houses in the USA and Canada and 55% in Europe and Australia had HVAC, but the diffusion rate of HVAC in Japan and Korea is low since houses are small. It may be that HVAC systems are not as effective as distributing sensors for small house.

Wren et al. proposed a method of detecting human movement that involves placing a lot of infrared sensors on the ceiling[5]. Sensors are placed in a lattice at intervals of a few meters, and events such as passing through rooms, changing direction, and passing each other are detected by analyzing the order of outputs from adjacent sensors. Hundreds of sensors are used, which leads to a high cost of installation and maintenance. Wilson et al. proposed a method to detect movements of multiple users by using infrared sensors and touch sensors on doors[6]. This system is realistic since relatively few sensors are required: one infrared sensor per room and one touch sensor per door. However, ordinary houses are not equipped with touch sensors, so these sensors would need to be installed for the purpose of movement detection. We aim to utilize sensors secondarily that are already installed in houses for other purposes. Therefore, we use only infrared sensors which are already set for lighting control purposes.

In addition, conventional studies employ probability-based methods such as those using Bayesian networks and particle filters[7]. These methods require a large amount of sensor data and ground truth, which must be collected through real living activity. Models of other houses do not works since the floor plans are different. In order to collect data for training models, one subject has to live alone for several days since it is hard to separate data if multiple subjects are living in the same house[6]. Collecting data after the house is constructed and before the owner moves in is possible, but it is not realistic to have a third party live in the home for a few days before the owner of the house moves in. Therefore, we aim to detect movements in houses without having to collect data for training.

3 Environment

This section describes the environment from which we collected data for the experiment and evaluation. Data collection took place in an experimental house located in the campus of the Japan Advanced Institute of Science and Technology (JAIST). The floor plan of the house is shown in Figure 1. The house is a two-story detached house consisting of 12 rooms and spaces: an entrance, living & dining room, lavatory, bathroom (in Japan, a room only for bathing), Japanese-style room, and WC (toilet) on the first floor, and two Western-style rooms, a bedroom, a spare room, and a WC on the second floor. In this paper, the lavatory and two WCs are treated as rooms. At least one sensor is installed in each room except for the bathroom since infrared sensors are not water-proof and are sensitive to heat. Apart from the rooms, sensors are placed in the hall, corridor, and stairway on the first floor, and in the corridor and stairway on the second floor. A total of 25 sensors are installed in the house as shown in Figure 2(a).

The infrared sensor used is model SN-MP13 manufactured by NTT Advanced Technology Corporation (Figure 2(b)); Table 1 lists its specifications. All sensors are connected to a server in the house via a sensor controller, as shown in Figure 3. The sensors capture infrared-rays every 500 msec and transmit the data to the sensor controller. “Found” is transmitted with time and sensor ID if a moving object emitting infrared-rays, in other words, radiating heat, is in the range of the sensor; otherwise, “Not Found” is transmitted. Then, the sensor controller carries the message from the sensor to the server if the new state (Found or Not Found) is different from the last state and continues for one second (two consective samples). Please note that states of “Found” and “Not Found” do not indicate the existence of a human, but a human moving under the sensor, and “Not Found” might be produced while an inhabitant stays under the sensor.

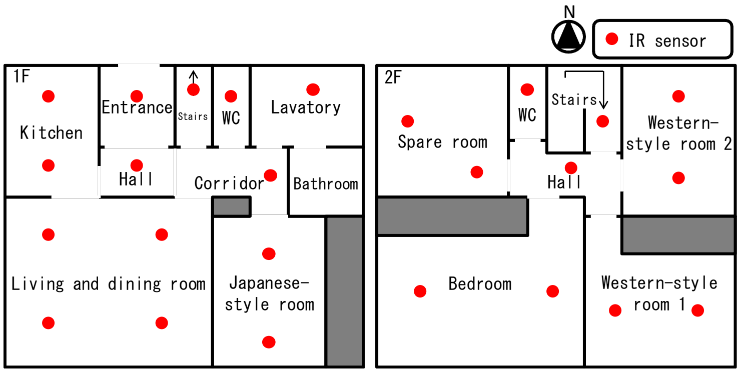
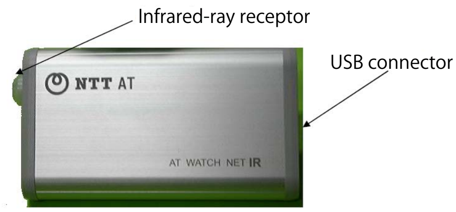


Fig. 1. Floor plan of experimental house and sensor arrangement. The shaded areas are closets and storage.

(a) Snapshot of sensor on the ceiling.



(b) Infrared sensor.

Fig. 2. Passive IR sensor.

Table 1. Sensor specifications.

Item		Value
Detection range		<5 m
Detection angle	Horizontal	38°
	Vertical	22°
Price		5,900 JPY (tax included)

Therefore, we do not use "Not Found" information since "Not Found" does not always mean that no one is there.

It might seem that tracking inhabitants is easy just by connecting sensors in the order of reaction time, but it is difficult for the following reasons. Infrared sensors exhibit characteristics in which they do not always react when the inhabitant walks along the edge of the detection area (false negative), and do sometimes react when the inhabitant walks near the doorway of the room (false

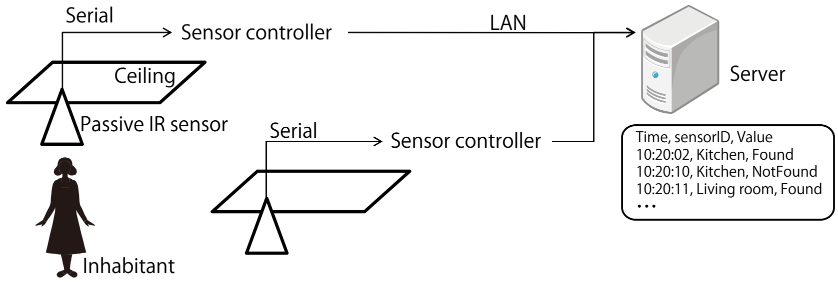


Fig. 3. Sensor configuration.

positive). Moreover, infrared sensors detect the existence of moving objects only and are not able to identify the person. “Found” information is mixed in the time-series when multiple inhabitants moved at the same time. For example, if a sequence of sensor IDs (s1, s2, s3, s4, s5) is produced by inhabitant A, and (s6, s7, s8, s9) is produced by inhabitant B, the server receives (s1, s2, s6, s7, s3, s6, s4, s9, s5). It is not clear who produced each ID. The distribution of sensors is not dense, and the distance between two sensors is long, which makes it difficult to trace a particular inhabitant by connecting sensor IDs on the floor map.

We propose here a method to detect movements from room to room even if multiple people move at the same time. For the sake of simplicity, we assume the inhabitants move between two different rooms, do not stop in the corridor, and do not stroll without a destination after leaving a room.

4 Movement detection method

This section describes the method to detect transition information from sensor sequences. In contrast to sensors that detect physical values such as hygrometers and accelerometers, infrared sensors output binary values. Physical values change according to the surrounding conditions, and changes in the environment or malfunctions are detected by analyzing the change in values. The physical value itself is preprocessed and then fed into a kind of classifier, such as a support vector machine (SVM)[8] and K-nearest neighbor (K-NN). On the contrary, a reacting sensor ID changes according to human movement. Therefore, we focus on the sensor ID only.

The procedure of the proposed method is described as follows.

1. Input ID sequence S.
2. Replace sensor IDs in S with a unique letter, i.e., sensor 1 is ‘a’, sensor 2 is ‘b’.
3. Calculate accordance ratio $A(i, t)$ of S_T with all templates, where S_T is a subset of S over a certain length of a window starting from time $t = T$, i is a sequential number of templates ($i = 1, \cdots, N$), and accordance ratio is a

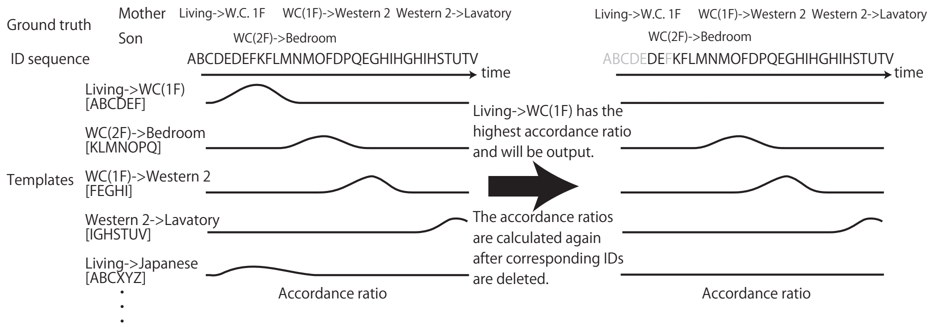


Fig. 4. Movement detection.

proportion of a template included in S_T. Iterate this calculation from t=0 to T_{end}, where T_{end} is the time of the last sample in S.
4. Search a pair (i', t') with the highest accordance ratio from A as follows;

$$(i', t') = \arg\max_{i,t} A(i, t). \tag{1}$$

If multiple pairs meet the above condition, the one with the longest template is selected. Moreover, if multiple pairs are still meeting the condition, the one with the earliest t is selected.
5. Quit if $A(i', t') < Threshold$, otherwise go to step 6.
6. Output the label of template $\#i'$, find the matching letters in S_T with template $\#i'$, and delete the matching letters from S. Then, go back to step 3.

The detailed flow in steps 3 and 4 is illustrated in Figure 4. Given the sequence of IDs produced by the mother's movements of *living room to WC(1F)*, *WC(1F) to Western-style room 2*, and *Western-style room 2 to lavatory*, and the son's movement *WC(2F) to bedroom*, the accordance ratio rises at the corresponding time for each template. THe transition *living to WC(1F)* is output since its accordance ratio is the highest, and the matching letters (A, B, C, D, E, F) are deleted from the original ID sequence, and the accordance ratio is calculated again. Deleting the corresponding letters will prevent the same output from being detected twice. Even when multiple inhabitants move, IDs derived from different movements would not be deleted, and all the transitions can be extracted.

A detailed algorithm to calculate the accordance ratio leverages a dynamic programming (DP) matching algorithm. By comparing m-length $S_T = (s_1, \cdots, s_m)$ and n-length template $Y = (y_1, \cdots, y_n)$, we can define a $m \times n$ matrix d_{ij}, where $d_{ij} = 0$ if $x_i = y_i$, and $d_{ij} = 1$ if $x_i \neq y_i$. Subsequently, path $P = (p_1, \cdots, p_k)$ is found, which is a pair of indices of S_T and Y. At this time, path P meets the following three conditions.

– Boundary condition
  $p_1 = (1, 1), p_k = (m, n)$

- Seriality
  $p_k = (a, b), p_{k-1} = (a', b') \Rightarrow a - a' \leq 1 \wedge b - b' \leq 1$
- Monotony
  $p_k = (a, b), p_{k-1} = (a', b') \Rightarrow a - a' \geq 0 \wedge b - b' \geq 0$

To find the path with the lowest cost that meets the above conditions, the following steps are applied.

1. **Initialization**:
   $Cost(0, 0) = 0$
   $Cost(i, 0) = \infty$ for $i = 1, \cdots, m$
   $Cost(0, j) = \infty$ for $j = 1, \cdots, n$
2. **Cost calculation**:
   Do for $i = 1, 2, \cdots, m$
   Do for $j = 1, 2, \cdots, n$

$$Cost(i, j) = min \begin{cases} Cost(i-1, j-1) \\ Cost(i-1, j\quad) \\ Cost(i\quad, j-1) + d(x_i, y_j) \end{cases}$$

3. **Lowest-cost path search**:
   $k = 0, i = m, j = n, p_k = (i, j)$
   While $i \neq 1 \&\& j \neq 1$:
   if $Cost(i-1, j-1) < Cost(i-1, j) \&\& Cost(i-1, j-1) < Cost(i, j-1)$
   $i--, j--, k++, p_k = (i, j)$
   else if $Cost(i-1, j) < Cost(i, j-1)$
   $i--, k++, p_k = (i, j)$
   else
   $j--, k++, p_k = (i, j)$

4. **Output**:
   Return P

Basically, Y is a subset of S_T when a transition is finished. Therefore, $d(x_i, y_j)$ is added to the second equation in order not to allow Y to grow even though the letters do not match. The number of matching letters c is counted from P, and the accordance ratio is obtained by c/n. For example, $Y = (a, b, c, d, e)$, and a, c, d, and e are matched to S_T in this order, and the accordance ratio becomes 0.8.

5 Evaluation

5.1 Data collection

The data used for the evaluation were collected from the scenarios from one and two subjects in the two-story house introduced in Section 3. In the one-subject scenario, the subject moved along a route that includes all possible transitions among 11 rooms excluding the bathroom. The transition from/to the living room was done in two ways since it has two routes: direct and via the kitchen. In the two-subject scenario, the subjects freely moved around the house for 10 minutes. Their transitions are listed in Table 2, which includes passing each other and staying together in the room.

Templates for training are made automatically by outputting all sensor IDs on a route with a given origin and destination. On routes that include a room with more than two sensors, all the combinations are included for the room. For an example of a transition from the living room to the bedroom, since the living room has four sensors that react in $2^4 - 1 = 15$ patterns and the bedroom has two sensors that react in $2^2 - 1 = 3$ patterns, the number of templates for the transition is $15 \times 3 = 45$. The size of window W is set to 25 seconds. This is because one transition takes up to 25 seconds in the data collection. We discuss the performance of our proposed method based on detection accuracy. The accuracy is measured on the basis of:

- If either the origin or destination of a detected transition is false, it is false positive.
- If both the origin and destination of a detected transition are correct, and
  - if the transition is detected in 25 seconds after the subject leaved the origin room, it is true positive.
  - Otherwise, it is false positive.

The departure and arrival times of the transition are recorded.

5.2 Result

The recall, precision, and F-measure of the results of movement detection for one and two subjects are plotted in Figures 5 and 6. Recall, precision, and F-measure are measured by $Recall = \frac{\#\ of\ true\ positive}{\#\ of\ test\ sample}$, $Precision = \frac{\#\ of\ true\ positive}{\#\ of\ detected\ sample}$, and $F-measure = \frac{2 \cdot Recall \cdot Precision}{Recall + Precision}$. The results indicate that recall increases as the threshold decreases since the number of outputs increases, while precision decreases also since uncertain results are output. On the contrary, precision increases as threshold increases since only reliable results are output, while recall decreases since the number of false negatives increases. The F-measure results show that the highest F-measures for one subject, 0.953, and for two subjects, 0.821, are obtained when the threshold is set to 0.8. The reason incorrect detections occur is that the sensor in the living room which is nearest to the kitchen incorrectly reacts when the subject goes to the kitchen. This is because the sensor is so close to the kitchen that the movement is detected as "to living room". In addition, some sensors did not work well because their position and direction were not very good. Therefore, we investigated which sensors had an important role.

Firstly, the F-measures for the cases when all 25 sensors were used and when one of the sensors in the non-room places was not used are shown in Figure 7. The bar chart of "w/o hall", for example, indicates the result when 24 sensors

Table 2. Scenario for two subjects

	Departure time	Origin	Arrival time	Destination
Subject A	15:46:54	Western-style room 2	15:47:05	Bedroom
	15:47:54	Bedroom	15:48:09	WC(2F)
	15:49:54	WC(2F)	15:50:13	WC(1F)
	15:50:54	WC(1F)	15:51:04	Lavatory
	15:51:34	Lavatory	15:51:46	Japanese-style room
	15:52:14	Japanese-style room	15:52:31	Entrance
	15:52:54	Entrance	15:53:07	Kitchen
	15:53:34	Kitchen	15:53:54	Western-style room 2
	15:54:29	Western-style room 2	15:54:36	Western-style room 1
	15:54:54	Western-style room 1	15:55:05	Spare room
	15:55:29	Spare room	15:55:52	Living room
	15:56:29	Living room	15:56:51	Bedroom
Subject B	15:47:54	Living room	15:48:05	Lavatory
	15:48:54	Lavatory	15:49:15	Bedroom
	15:50:24	Bedroom	15:50:34	Spare room
	15:51:24	Spare room	15:51:36	Western-style room 2
	15:52:24	Western-style room 2	15:52:34	WC(2F)
	15:53:24	WC(2F)	15:53:36	Western-style room 2
	15:54:24	Western-style room 2	15:54:39	Living room
	15:55:24	Living room	15:55:31	Kitchen
	15:56:24	Kitchen	15:56:36	Japanese-style room

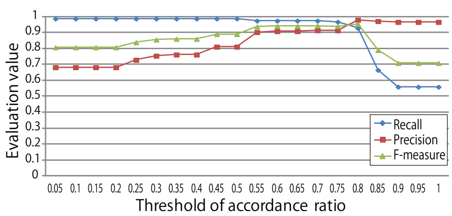


Fig. 5. Result of one subject.

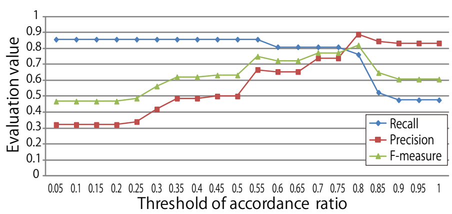


Fig. 6. Result of two subjects.

were used, excluding the sensor in the hall on the first floor. The use of all sensors resulted in an F-measure of 0.95, while the F-measure is dropped to 0.92 when a sensor in the hall on the first floor was excluded. This is the lowest performance of the five places. The same trend was confirmed for two subjects. This means that the sensor in the hall on the first floor plays an important role. This is because this sensor is involved in many patterns of transitions.

On the contrary, the F-measure rose up to 0.98 when the sensor in the hall on the second floor for one subject. This is because the sensor setting was not adequate, and it seemed to be a malfunctioning sensor. The results of the stairways indicate that the performance does not drop when the sensors in the stairways were excluded. This is because the sensors complementarily react to each other.

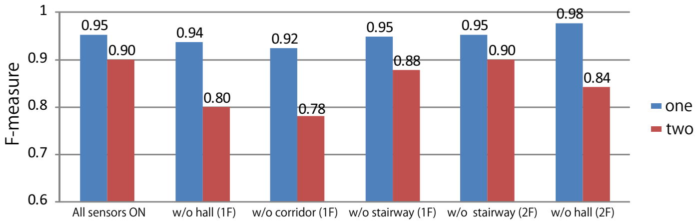


Fig. 7. Results for two subjects.

However, the use of both sensors would be important when the two subjects passing each other on the stairs. These results indicate that it is important to place sensors at points where inhabitants pass each other.

Finally, the F-measures for the case when only one of multiple sensors installed in each room was used is shown in Figure 8. The bar chart of "Living East/Far," for example, is the result when one out of four sensors on the east side and far from the doorway, and 21 sensors outside the living room were used. "Far" and "Near" denote the sides far from the doorway and near the doorway. The results of the living room show that the performance of the sensor at West/Near was the lowest of the four sensors in the living room. This is because the West/Near sensor is close to the kitchen, which incorrectly reacted just when the subject entered the kitchen. However, the F-measure of living West/Far dropped to 0.76 for two subjects. This is because the subjects did not go to the far side of the living room. The performance of sensors at the far side in the kitchen, Japanese-style room, bedroom, Western-style room 1, Western-style room 2, and spare room was also lower than that of sensors at the near side. Accordingly, a sensor placed at the near side performed better than that at the far side since the inhabitant does not always go to the far side of the room. These results indicate that it is more effective to place sensors at the near side of the doorway. However, sensors placed very close to the doorway may output a false positive just when the inhabitant walks through the front of the room. Therefore, sensors should be placed taking their detection range and angle into consideration.

6 Conclusion

We proposed a method to detect room-to-room transitions in a home environment by using only some infrared sensors. An evaluation was carried out with the data collected in an experimental two-story house for one and two test subjects, and an F-measure of 0.82 was confirmed, which is a relatively high degree of performance even though relatively few sensors were used, and our proposed method does not require data collection for training. In addition, we investigated the performance when changing the combination of sensors and clarified that sensors place where transitions cross contribute to improving accuracy and

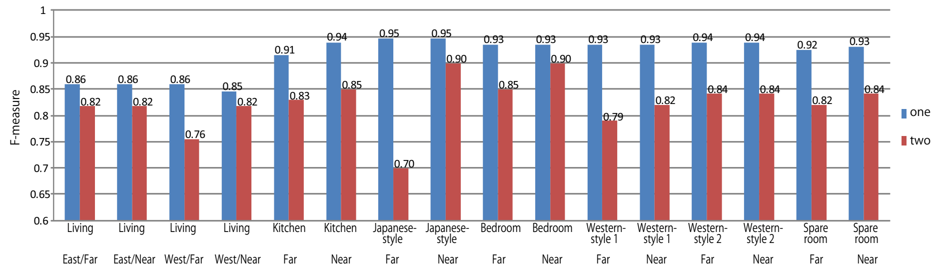


Fig. 8. Results for two subjects.

that removing malfunctioning sensors also improves accuracy. Sensors in rooms should be placed near the doorway taking into consideration their detection range and angle.

We are working on a method to identify people and to connect transitions, and will conduct further investigations on the sensor arrangement considering their detection area, angle, and placement.

References

1. Wren, C., Azarbayejani, A., Darrell, T., and Pentland, A.: Pfinder: Real-time Tracking of the Human Body, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 19, No. 7, pp. 780–785 (1997).
2. Yamahara, H., Takada, H., and Shimakawa, H.: Behavior Detection Based on Touched Objects with Dynamic Threshold Determination Model, in *Proc. of the 2nd European Conference on Smart Sensing and Context (EuroSSC 2007),* pp. 142–158 (2007).
3. Kamisaka, D., Muramatsu, S., Iwamoto, T., and Yokoyama, H.: Design and Implementation of Pedestrian Dead Reckoning System on a Mobile Phone, *IEICE Transactions,* Vol. E94.D, No.6, pp. 1137–1146 (2011).
4. Patel, S. N., Reynolds, M. S., and Abowd, G. D.: Detecting Human Movement by Differential Air Pressure Sensing in HVAC System Ductwork: An Exploration in Infrastructure Mediated Sensing, in *Proc. of International Conference on Pervasive Computing (Pervasive 2008),* pp. 1–18 (2008).
5. Wren, C. R. and Munguia-Tapia, E.: Toward Scalable Activity Recognition for Sensor Networks, in *Proc. of the International Workshop in Location and Context-Awareness (LoCA 2006),* pp. 168–185 (2006).
6. Wilson, D. H. and Atkeson, C. G.: Simultaneous Tracking and Activity Recognition (STAR) Using Many Anonymous, Binary Sensors, in *Proc. of International Conference on Pervasive Computing (Pervasive 2005),* pp. 62–79 (2005).
7. Doucet, A., Freitas de, N., and Gordon, N.: Sequential Monte Carlo Methods in Practice, Springer (2001).
8. Vapnik, V.: The Nature of Statistical Learning Theory, Springer (1995).