# Person Identification Based on the Matching of Foot Strike Timings Obtained by LRFs and a Smartphone

Kenji Koide and Jun Miura
*Department of Computer Science and Engineering, Toyohashi University of Technology, Japan*

*Abstract*— **This paper describes a person identification method using a smartphone and laser range finders (LRFs) for a mobile service robot. The robot is equipped with LRFs and the target person holds a smartphone. The method first detects the foot strike timings of the target person using the smartphone and those of all people by using the LRFs. By finding the person whose foot strike timings captured by the LRFs are similar to those obtained by the smartphone, the robot can identify the target person. Person identification experiments and person following experiments are conducted in order to validate the method. Since the method only requires a person to simply hold a smartphone, it can be easily applied to daily situations.**

## I. Introduction

Person identification is one of the fundamental functions of mobile service robots which work in everyday environments. Such robots should be able to follow the target person and provide appropriate services. Since environments in daily life are sometimes heavily populated, the robot may lose the track of the target person. In that case, the robot has to be able to re-identify the target person in order to continue to provide services.

Many previous works use a combination of laser range finders (LRFs) and cameras for person tracking and identification by mobile robots [1], [2], [3]. In these works, a robot learns the appearance of the target person from cameras and uses them for identifying the person. However, if the appearance of the target person changes drastically while the person is occluded by other persons or obstacles, it is difficult for the robot to find the correct person again using only the image. If we use a device for identification and the target person holds the device, we can realize a person identification which does not suffer from any environmental changes and persistent occlusions of the target person.

Some works propose person identification methods using environment sensors and an IMU (Inertial Measurement Unit) [4], [5]. They place multiple static environment sensors, such as a camera and an LRF, in an environment and attach an IMU to the target person. These methods measure the walking pattern of the target person using the IMU and those of all persons in the environment using the environment sensors. Shiomi et al. [4] use depth cameras as environment sensors. They calculate a persons' acceleration using the depth cameras and the IMU, and classify the persons' states into moving or stopping. By matching the states obtained by the depth cameras and the IMU, they can identify the target person. Ikeda et al. [5] put LRFs in an environment, and tracked the legs of all the persons in the environment. They estimate the acceleration of the legs and compare them with the acceleration obtained by the IMU. By calculating the signal correlation between the accelerations, they find the person holding the IMU among others. However, we cannot apply these methods to mobile robots directly since multiple static sensors are not available.

We proposed a person identification method based on the matching of foot strike timings for mobile robots [6]. In the method, an IMU is put on a foot of the target person and the robot is equipped with two LRFs. The proposed method estimates the foot strike timings of the target person using the IMU and those of all persons using the LRF. By matching these data, the robot can reliably identify the target person in the LRF data. However, since the IMU is put on a foot, the method can detect and use only the strike timings of one foot. In addition to that, the method requires the person to hold the unusual device.

In this paper, we extend our previous method to being able to use a smartphone in a pocket. By use of the smartphone, the method becomes applicable in daily situations. Moreover, the method can detect the strike timings of both legs, and classification performance is improved compared with the previous method.

The remainder of this paper is organized as follows. Sec. II explains an overview of the proposed person identification system. Sec. III and Sec. IV describe the methods for measurement of foot strike timings and stopping states of a person using LRFs and a smartphone, respectively. Sec. V describes a method of integrating both sets of foot strike timing data for person identification. Sec. VI shows experimental results and the comparison with our previous method. Sec. VII concludes the paper and discusses future work.

## II. System Overview

Fig.1 shows an overview of the proposed system. The robot is equipped with two LRFs, and the target person puts a smartphone in their pocket. The smartphone is connected to the robot via wifi. The system first detects and tracks all persons around the robot using the top LRF, and then estimates foot strike timings for each person using the bottom LRF. It also estimates the timings from the acceleration of the smartphone, and compares it with the timings of all tracked persons to identify the target person. When the target person is stopping, however, the foot strike timing is not available. We therefore judge if a person is stopping by using the LRFs and the smartphone, and the stopping states are also

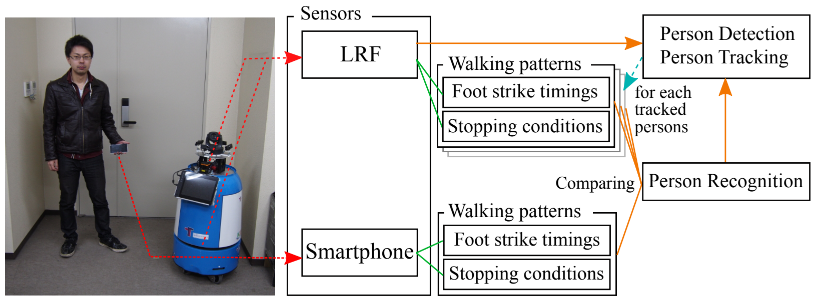



Fig. 1. System overview.

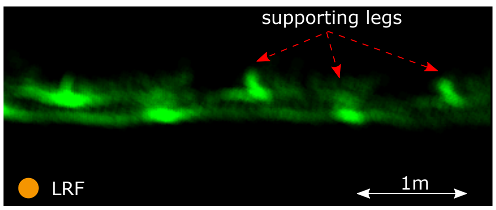



Fig. 2. Accumulated range data of the legs of a walking person.

compared to estimate the target person. Note that the top LRF is used just to simplify the data association for people tracking. Therefore, this method essentially requires only one LRF placed at a leg height.

We define a dissimilarity measure for the timings and the stopping states from the LRFs and the smartphone, and calculate the likelihood that each person is the target. This information is integrated over time using a Bayesian inference, and the person with the highest posterior probability is judged as the target. When the LRF-based tracker loses the target person due to, for example, occlusion, it stops the estimation and starts reacquiring the target person among all persons.

## III. Estimation of Foot Strike Timings and Stopping State using LRFs

For mobile robots, we developed a method for estimating foot strike timings using LRFs [6]. The method first detects the positions of the supporting legs of a walking person from LRF data and then estimates the strike timing from a time period where a foot is near each supporting leg position.

Fig. 2 shows an example of accumulated range data of a walking person obtained from the LRF placed at a leg height. The supporting legs of the person appear as high-density regions of range data. According to Nakamura et al. [7], we extract the supporting legs by spotting high-density regions using Mean Shift [8]. We then estimate the actual positions of the supporting legs using a maximum likelihood estimation which takes the self-occlusions of the supporting legs into account.

Let $X = [x_1, y_1, \cdots, x_n, y_n]$ be the positions of the supporting legs, $Y = [x'_1, y'_1, \cdots, x'_n, y'_n]$ be their observed positions, and $\Sigma = [\sigma_1^2, \cdots, \sigma_n^2]$ be the observation variances. The Likelihood function $L$ is defined as:

$$L = \prod_{i=1}^{n} \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{(x'_i - x_i)^2 + (y'_i - y_i)^2}{2\sigma_i^2}\right) \quad (1)$$

We minimize the following objective function $J$.

$$\begin{aligned} J &= -\log L \\ &= \sum_{i=1}^{n} \log 2\pi\sigma_i^2 + \sum_{i=1}^{n} \frac{1}{2\sigma_i^2} \left\{(x'_i - x_i)^2 + (y'_i - y_i)^2\right\} \end{aligned} \quad (2)$$

Assuming a steady gait (the step length is constant), we obtain the following constraint function:

$$\begin{aligned} (x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 = \text{const.} \\ i = (1, 2, \cdots, n-1, n) \end{aligned} \quad (3)$$

From this equation, we obtain:

$$\begin{aligned} (x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 - \\ (x_{i+1} - x_i)^2 - (y_{i+1} - y_i)^2 = 0 \\ i = (2, 3, \cdots, n-2, n-1) \end{aligned} \quad (4)$$

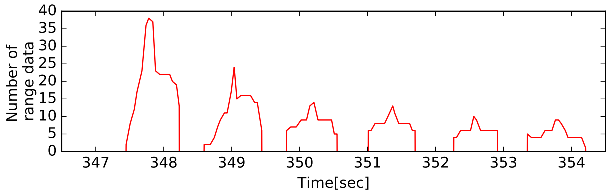
We minimize eq. (2) subject to eq. (4) using the method of Lagrange multiplier to obtain the estimation of supporting leg positions.

We use five steps for the estimation of supporting leg positions. We assume that the walking speed is constant throughout the duration.

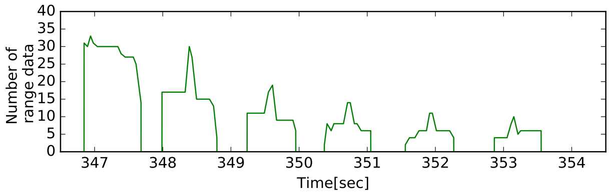
When the robot observes a walking person from the side, a leg on the robot side is always visible while the other is sometimes occluded. We thus give the observation of the supporting leg on the robot side a small variance (i.e., high reliability) and that of the other leg a large variance (low reliability).

To estimate foot strike timings from the positions of the supporting legs, we count the number of the range data around a supporting leg at each frame, and examine how the number changes over those frames. Fig. 3 (a)(b) show the change of the numbers of range data for both legs. Each cluster corresponds to a foot strike. We treat the number as a weight and consider the weighted mean of times as a strike timing. Fig. 3 (c) shows the estimated foot strike timings.

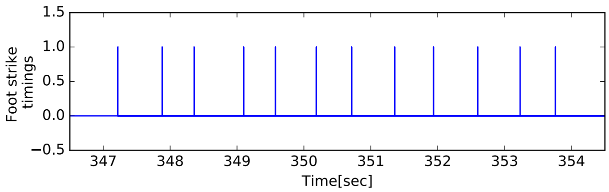
The stopping state of a person is determined by the walking speed measured by the LRF-based tracking. If the speed of a person is less than a specified threshold (currently, 0.3 [m/sec]) the person is considered to be stopping.




(a) Change of the number of range data for one leg.




(b) Change of the number of range data for the other leg.




(c) Estimated foot strike timings.

Fig. 3. Estimating foot strike timings using LRFs.

## IV. Estimation of Foot Strike Timings and Stopping State using a Smartphone

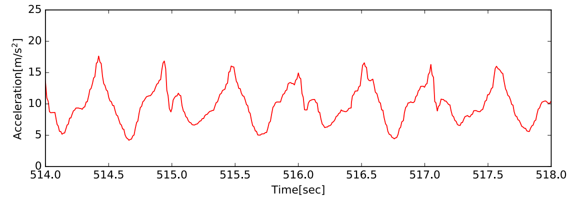
While a person is walking, the body of the person moves up and down periodically. By detecting the peak of the acceleration of the body, we can detect the foot strike timings. We use a method proposed by Li et al. [9]. They first apply an FIR low-pass filter to the acceleration data obtained by a smartphone, and then detect the peak of the filtered acceleration using two threshold values $\Delta t$ and $\Delta a$ (see Fig. 4(b)). Since their method does not depend on the position of the sensor, the smartphone can be held at any location on a person; the person can put a smartphone in a pocket or hold it in the hand. Fig. 4 shows an example of the acceleration of a smartphone placed in the chest pocket and the estimated foot strike timings from it. We set the cutoff frequency of the low-pass filter 3 [Hz], $\Delta t = 0.2$ [sec], and $\Delta a = 1.5$ [m/s$^2$].

The stopping state of a person is determined when the smartphone is judged as being stationary for a certain period of time. We use Jimenez's method [10], which uses simple thresholds of acceleration and angular velocity to judge if a sensor is stationary or not. If the smartphone is judged as stationary for 1.0[sec], we consider that the person who has the smartphone has stopped.

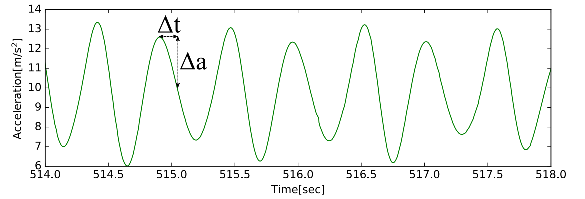
Fig. 5 shows an example of foot strike timings and stopping states when a person walks for several seconds and then stops and walks again.

## V. Data Integration for Person Identification

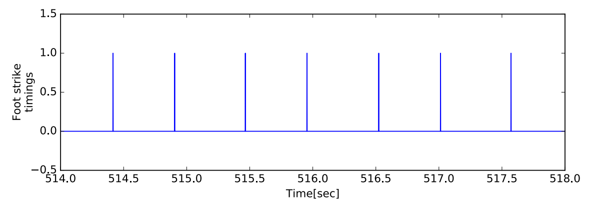
Stopping states or foot strike timings are obtained by a smartphone for the target person and by LRFs for all persons. We compare each person's data with those of the target to find the best-matched person in the LRF data. For this




(a) Raw acceleration.




(b) Low-pass filtered acceleration.




(c) Estimated foot strike timings.

Fig. 4. Estimating foot strike timings using a smartphone.

TABLE I
DISSIMILARITY OF FOOT STRIKE TIMINGS AND STOPPING STATE

| | target person | | other persons | |
|---|---|---|---|---|
| | foot strike | stopping | foot strike | stopping |
| # of data | 381 | 472 | 1483 | 2486 |
| mean | 63.6 [msec] | 0.0585 | 144.8 [msec] | 0.3005 |
| std. dev. | 28.6 [msec] | 0.0898 | 68.2 [msec] | 0.2650 |

purpose, we define the dissimilarity of the stopping states or the timings, which is then used for defining the likelihood function. The likelihood function is used for applying the Bayesian estimation to the target identification.

### *A. Dissimilarity Measure between LRF and Smartphone Data*

To compute a dissimilarity between foot strike timings, we first associate the timings by LRF and those by a smartphone by finding the closest LRF timing for each smartphone timing. We calculate the mean of the time differences between the associated timings, and use it as the dissimilarity measure.

The dissimilarity between stopping states by LRF and smartphone is calculated by the difference between the binary patterns of stopping. We set a time window and measure the total time duration where the LRF and the smartphone patterns are different. The duration is normalized by the width of the time window and used as the dissimilarity measure. We set the time window size to 5 [sec].

Table I gives statistics of dissimilarity values for foot strike timings and stopping states. This is obtained from the real experiments which were conducted under the same settings used in Sec. VI. The dissimilarities of the target person is much smaller than the dissimilarities of the others.

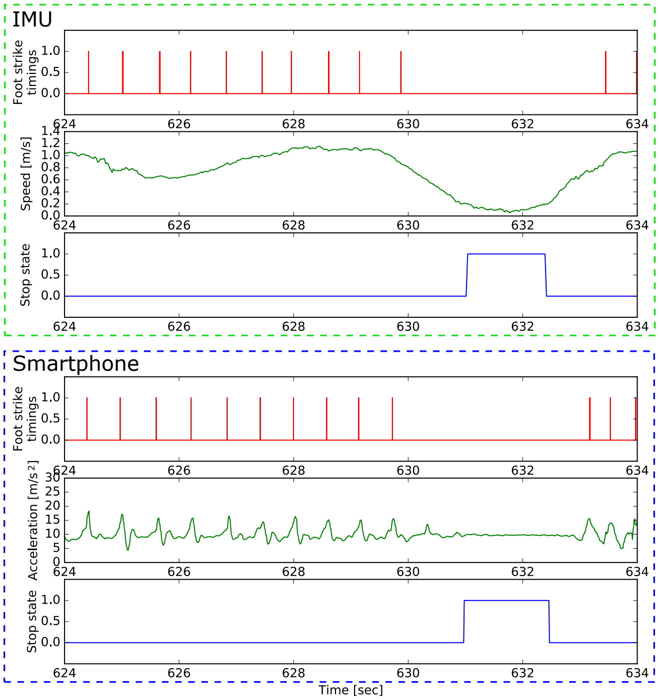



Fig. 5. Foot strike timings and stopping states obtained by LRFs and a smartphone.

### B. Bayesian Estimation for Person Identification

The target person usually shows low dissimilarities and the others high dissimilarities. The target person, however, sometimes shows a high dissimilarity due to a lack of range data or measurement errors. For a robust tracking, we use a Bayesian estimation for determining the target person.

Let $p(x_i)$ be the prior probability that the $i$th person in the LRF data is the target person. We define the likelihood of person $x_i$ for an observed dissimilarity value $y$ as:

$$p(y|x_i) = \exp\left(-cy\right) \tag{5}$$

where $c$ is a constant. We calculate the likelihood values for the foot strike timing and the stopping state, and the multiplication of the two likelihood values is used as the likelihood $p(\boldsymbol{y}|x_i)$. Then the posterior probability $p(x_i|\boldsymbol{y})$ is given by:

$$p(x_i|\boldsymbol{y}) = \alpha p(\boldsymbol{y}|x_i)p(x_i), \tag{6}$$

where $\alpha$ is the normalization constant. The probability is updated every 100 [msec].

### C. Re-detection of the Target Person

If the LRF-based tracking loses the target person, Bayesian estimation stops temporarily and the system starts to re-detect the person. This re-detection is done while the target person is walking, by searching for a person whose foot strike timings by LRF are close enough with those by a smartphone to a high confidence.

We consider that a pair of foot strike timings matches if their difference is less than a threshold $th_{tm}$. Then, if the foot strike timings of a person (by LRFs) and those of the target person (by a smartphone) have at least $n_{sim}$ matched frames in $n_{test}$ consecutive frames, that person is considered as the target. We determine these three parameters as follows.

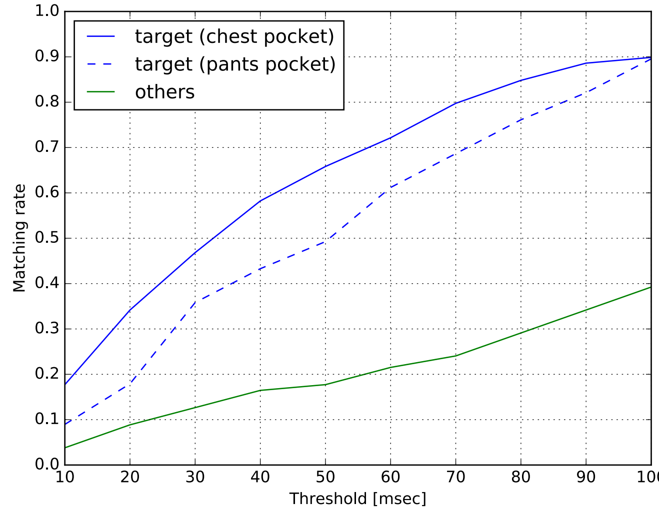



Fig. 6. Matching rate.

Fig. 6 shows relationship between the matching threshold $th_{tm}$ and the matching rate of timings for the actual target person and other persons. We calculated the matching rate from a real data sequence. The experimental setting is the same as the one used in Sec. VI-A. In the experiments, we placed a smartphone in two difference locations, a trousers pocket and a chest pocket. As the threshold increases, the rates increase, but that for the target person much more rapidly increases. When we put the smartphone in the trousers pocket, the matching rate of the target person is less than the chest pocket case. It is strongly affected by the attached leg and the foot strike of the opposite leg becomes difficult to detect. Even in the case of the trousers pocket, however, since the matching rate of the target person is significantly larger than the others, it can be used for identifying the target person.

According to binomial distribution, we can calculate the probability that a person is identified as the target from the matching rate and arbitrarily $n_{sim}$ and $n_{test}$ as:

$$p_{ident}(p, n_{sim}, n_{test}) = \sum_{i=n_{sim}}^{n_{test}} \binom{n_{test}}{i} p^i (1-p)^{n_{test}-i} \tag{7}$$

where $p$ is the matching rate of the foot strike timings of the person. We calculate the probability where the correct person is identified as the target (true positive rate) and where another person is identified as the target (false positive rate).

In order to determine the appropriate parameters, we set the target true positive rate as 80 % and the target false positive rate as 5 %. Table II shows examples of the candidate parameters which meet the criteria. As shown in Table II, if we set $n_{sim}$ and $n_{test}$ large enough (it means focusing on the person for a longer number of seconds), we can obtain better re-detection performance.

For mobile robots, however, there is a limitation on re-detection time, because the target person may get far away from the robot while the robot is trying to re-detect the person. We, therefore, choose parameters $n_{sim} = 3$ and

TABLE II

CANDIDATE PARAMETERS

| threshold | $n_{sim}$ | $n_{test}$ | true positive | false positive |
|---|---|---|---|---|
| 70 [msec] | 3 | 4 | 0.814 | 0.045 |
| 70 [msec] | 4 | 5 | 0.831 | 0.027 |
| 90 [msec] | 4 | 5 | 0.897 | 0.049 |
| 90 [msec] | 5 | 6 | 0.857 | 0.019 |
| 90 [msec] | 6 | 7 | 0.814 | 0.007 |
| 70 [msec] | 6 | 9 | 0.910 | 0.008 |

TABLE III

RESULTS OF THE PERSON IDENTIFICATION EXPERIMENT

| duration [sec] | | smartphone | | IMU |
|---|---|---|---|---|
| | | chest | trousers | foot |
| successfully tracked | | 135.5 | 128.2 | 121.8 |
| lost track | target appears | 44.6 | 49.1 | 56.6 |
| | target not appears | 28.3 | 24.5 | 27.7 |
| tracked wrong person | | 0.0 | 6.6 | 2.3 |
| average identification time | | 3.4 | 3.5 | 4.7 |

$n_{test} = 4$; then, the true positive rate and the false positive rate are estimated to be $0.814$ and $0.045$ respectively, and the minimum time for re-detection will be about $1.5$ [sec].

### D. Comparison with the previous method

We compare the identification performances of the proposed method and our previous method [6]. We define the identification performance as the harmonic mean of the true positive rate and the inverse of the false positive rate:

$$\text{performance} = \frac{2TP \cdot (1 - FP)}{TP + (1 - FP)} \tag{8}$$

and we define the identification time as the average of the minimum identification time ($n_{sim}$) and the maximum identification time ($n_{test}$):

$$\text{identification time} = T_{step} \frac{n_{sim} + n_{test}}{2} \tag{9}$$

where $T_{step}$ is the cycle time of foot strike. Since the proposed method uses the foot strike timings of both legs and the previous method uses only one leg, we set $T_{step}$ of the proposed method to $0.5$ [s] and the previous method to $1.0$ [s].

Fig. 7 shows the relationship between identification time and identification performance. The identification performance values shown in Fig. 7 are the best ones among those with the same identification time. As shown in Fig. 7, if we take a longer identification time, we can obtain better performance. Since the proposed method can obtain more foot strike timings than the previous method in a certain period, it shows a better identification performance than the previous method.

## VI. EXPERIMENTAL RESULTS

### A. Person Identification Experiment

We conducted a person identification experiment. In the experiment, the target person had two smartphones (FREETEL FTJ152B, MediaTek accelerometer is embedded), one in his trousers pocket and one in his chest pocket and attached an IMU (ZMP IMU-Z2) to his foot. We used two LRFs (HOKUYO UST-20LX) to track people and estimate

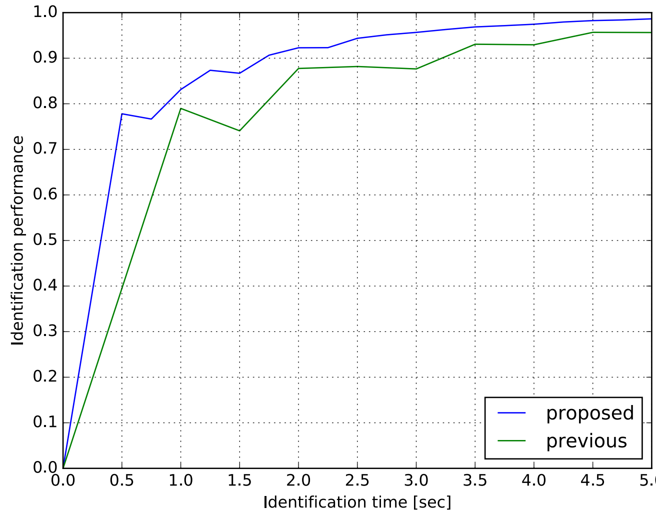
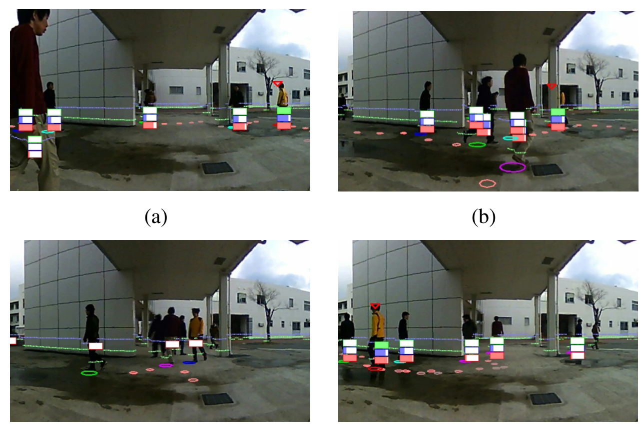



Fig. 7. Relationship between the identification time and the identification performance.



(a) (b) (c) (d)

Fig. 8. Person identification experiment. The person wearing the yellow jacket is the target person; he placed two smartphones, one in his chest pocket and one in his trousers pocket and an IMU on his foot. Circles under persons indicate their positions obtained by LRF-based tracking, red triangles indicate the target person and green, blue, and red bars indicate the posterior probabilities of each person, the likelihood of the foot strike timings, and the likelihood of the stopping state, respectively.

their foot strike timings. We apply the proposed method to those smartphones and our previous method to the IMU.

Fig. 8 shows snapshots of the experiment with the smartphone in the chest pocket. The person wearing the yellow jacket is the target person. Circles under persons indicate their estimated positions obtained by the LRF-based tracking. Green, blue, and red bars indicate the posterior probabilities that a person is the actual target, the likelihood of the foot strike timings, and the likelihood of the stopping state, respectively. Red triangles on a person indicate that the person is tracked as the target.

The target person and the other persons entered the field and walked about (Fig. 8 (a)), and then the target person was lost by the LRF-based tracking due to occlusions (Fig. 8 (b)(c)). Once the target person appeared again and walked several steps, however, the target person was re-detected and the tracking was resumed (Fig. 8 (d)). During the experiment, the robot lost the track of the target a total of 12 times due to occlusions. The target person was, however, successfully re-detected in every case.

Table III shows the result of the experiment. The total time of the experiment was 208.4 [sec]. In the case of the smartphone in the trousers pocket, while the target person is hidden by other persons, wrong persons are re-detected as the target twice. However, the system realized that the person is not the target by Bayesian inference and soon tracked the target person again. Since we used parameters which are optimized for a smartphone in a chest pocket, the identification performance in the case of the trousers pocket is a little worse than in the case of the chest pocket. If we optimize the parameters for a smartphone in a trousers pocket, we could improve identification performance.

Since the previous method takes longer identification time than the proposed method, its duration of target loss is longer than the others. It shows that the proposed method has improved its responsiveness from the previous method.

During the experiment, the whole procedure except visualization took less than 1 [msec] per frame. Since the processing cost is very low, the method can be extended to track multiple targets by using a smartphone of every target person.

### *B. Person Following Experiment*

We applied the person identification method to a person following task. While the robot is tracking the target person, it moves toward the person. If the track of the person is lost, the robot moves toward the target person position predicted from the latest observed position and the velocity of the person in order to keep the robot close to the person. We used a path planning method by Ardiyanto and Miura [11] to make the robot avoid obstacles while following the person.

Fig. 9 shows snapshots of a person following experiment. During the experiment, the robot lost the target person several times due to occlusion by others (Fig. 9 (b)). However, once the target person appeared and walked by several steps, the robot successfully re-detected him among the others and continued following him (Fig. 9 (c)).

(a)

(b)

(c)

Fig. 9. Specific person following experiment: The left column shows experiment scenes and the right column shows view of the camera on the robot. The meanings of the markers (triangle, circle, bars) are the same as in Fig. 8.

## VII. Conclusions and Future Work

This paper has described a method of identifying a specific person using LRFs and a smartphone. In the method, the robot is equipped with LRFs and the target person holds a smartphone. Both sensors are independently used for classifying the stopping states and for obtaining foot strike timings. The obtained stopping states and timings are integrated using a Bayesian inference to identify the target person. Since the method requires the person to just hold a smartphone, it can easily be applied to everyday situations. The method was validated through the person identification experiments and the person following experiments.

Smartphones provide plenty of sensor data which contains IMU, GPS, illuminance, and wifi signal strength. We are planning to extend the method to be able to utilize such information for more robust person identification.

## References


[1] N. Bellotto and H. Hu, "Multisensor data fusion for joint people tracking and identification with a service robot," in *Proc. of IEEE Int. Conf. on Robotics and Biomimetics*. IEEE, 2007, pp. 1494–1499.

[2] M. Luber, L. Spinello, and K. O. Arras, "People tracking in RGB-d data with on-line boosted target models," in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. IEEE, 2011, pp. 3844–3849.

[3] J. Satake, M. Chiba, and J. Miura, "A SIFT-based person identification using a distance-dependent appearance model for a person following robot," in *Proc. of IEEE Int. Conf. on Robotics and Biomimetics*. IEEE, 2012, pp. 962–967.

[4] M. Shiomi and N. Hagita, "Finding a person with a wearable acceleration sensor using a 3D position tracking system in daily environments," *Advanced Robotics*, vol. 29, no. 23, pp. 1563–1574, 2015.

[5] 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," *Ambient Intell Human Computing*, vol. 5, no. 5, pp. 645–654, 2013.

[6] K. Koide and J. Miura, "Person identification based on the matching of foot strike timings obtained by IMU and LRF," in *Proc. of IROS Workshop on Assistance and Service Robotics in a Human Environment*. IEEE, 2015.

[7] K. Nakamura, H. Zhao, X. Shao, and R. Shibasaki, *Human Sensing in Crowd Using Laser Scanners*. InTech, 2012.

[8] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 8, pp. 790–799, 1995.

[9] F. Li, C. Zhao, G. Ding, J. Gong, C. Liu, and F. Zhao, "A reliable and accurate indoor localization method using phone inertial sensors," in *Proc. of ACM Conf. on Ubiquitous Computing*. ACM, 2012.

[10] A. Jimenez, F. Seco, J. Prieto, and J. Guevara, "Indoor pedestrian navigation using an INS/EKF framework for yaw drift reduction and a foot-mounted IMU," in *7th Workshop on Positioning, Navigation and Communication*. IEEE, 2010, pp. 135–143.

[11] I. Ardiyanto and J. Miura, "Real-time navigation using randomized kinodynamic planning with arrival time field," *Robotics and Autonomous Systems*, vol. 60, no. 12, pp. 1579–1591, dec 2012.