Proc. 13th Int. Conf. on
Intelligent Autonomous Systems,
Padova, Italy, July 2014.

Specific Person Detection and Tracking by a Mobile Robot using 3D LIDAR and ESPAR Antenna

Kazuki Misu and Jun Miura

Department of Computer Science and Engineering
Toyohashi University of Technology

Abstract. Tracking a specific person is one of the important tasks of mobile service robots. This paper describes a novel and reliable strategy for detecting and tracking a specific person by a mobile robot in outdoor environments. The robot uses 3D LIDARs for person detection, and identification and a directivity-variable antenna (called *ESPAR antenna*) for locating a specific person even under occluded and/or out-of-view situations. A sensor fusion approach realizes a reliable tracking of a specific person in actual outdoor environments. Experimental results show the effectiveness of the proposed strategy.

Keywords: Person detection and tracking, Mobile robot, 3D LIDAR, ESPAR antenna

1 Introduction

Person detection and tracking is one of the fundamental functions of personal service robots. Such a robot should be able to avoid, follow, or communicate with persons and, therefore, knowing the person location is important. Person identification is also necessary for providing person-specific services such as attending and guiding.

Vision has been a primary sensor for person detection for robots and surveillance purposes. Various visual cues have been used such as appearance (e.g., local features like HOG [1]), stereo data [2], or motion (e.g., optical flow) [3]. Many works try to integrate multiple features for increasing the reliability [4]. Although vision can provide rich information, it is inherently sensitive to illumination condition.

Range sensors such as a LIDAR (laser imaging detection and ranging) have often been used for mobile robots to detect objects in the scene. They are also used for detecting people [5–8] but a limited number of scans are sometimes not enough for reliably detecting persons. High definition range sensors have recently been developed and shown to be effective for person detection in real outdoor environments [9–11].

Concerning identifying a person, simple color- and/or texture-based methods (e.g., [12]) are often used but have relatively low discriminative power. More

powerful vision-based methods such as face recognition (e.g., [13]) and gate identification (e.g., [14]) are popular but may be difficult to apply to bad illumination conditions and/or crowded situations where occlusions frequently occur. RFID tags can be used [15] but relatively large antennas should be equipped on a robot for detecting people in various directions.

This paper describes a method of detecting and tracking a specific person using 3D LIDARs and ESPAR (Electronically-Steerable Passive Array Radiator) antennas [16]. The ESPAR antenna can change its directivity very fast and is suitable for detecting the direction of a radio transmitter in a long range. We mainly use the 3D LIDAR for person detection and the ESPAR antenna for identifying a specific person with a radio transmitter but further investigate sensor fusion techniques for reliably continuing a tracking task under difficult situations such as long and full occlusions or out-of-view situations.

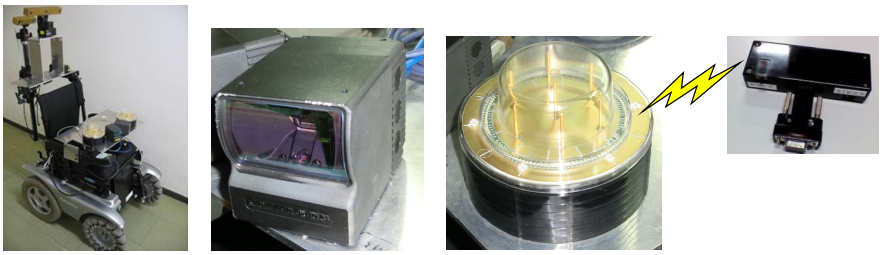
The contributions of the paper are: (1) the use of a new type of sensor (ESPAR antenna) for identification and localization, (2) a sensor fusion method for the two sensors for a robust specific person detection and tracking, and (3) its implementation and testing on an actual mobile robot.

The rest of the paper is organized as follows. Sec. 2 describes an overview of our robot system and the proposed method and that of the developed system. Sec. 3 briefly explains a 3D LIDAR-based person detection. Sec. 4 describes a person tracking using the 3D LIDAR and the ESPAR antenna. Sec. 5 describes the identification and localization of the target person using the ESPAR antenna. Sec. 6 shows experimental results of person tracking using the developed system. Sec. 7 summarizes the paper and discusses future work.

2 Overview of the System and the Proposed Method

2.1 Robot and sensors

Robot Fig. 1(a) shows our robot. The mobile base is an electric wheelchair with a four-wheel drive mechanism (PatraFour by Toyota Motor East Japan, Inc). The wheelchair is modified by the manufacturer so that it can be controlled by a PC.



(a) Our robot. (b) FX8 3D LIDAR. (c) ESPAR antenna (receiver) and transmitter.

Fig. 1. The robot and the sensors. Two LIDARs and two ESPAR antennas are put on the front part of the robot. Stereo cameras on the rear are not used in this research.

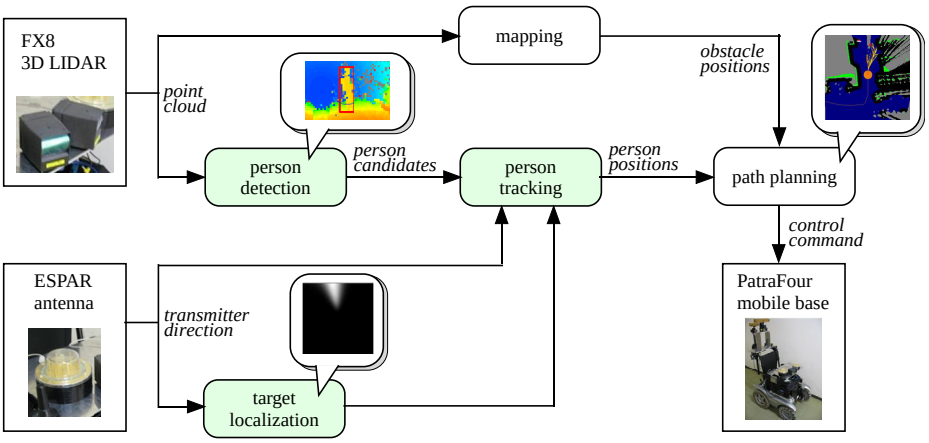


Fig. 2. Software configuration. Green-shaded blocks are new modules developed in this research.

Table 1. Specifications of 3D LIDAR (FX8).

Frame rate	4 ± 0.4 frames/sec.
Field of view	Horizontal 60° / Vertical 50°
Image size	97×61 (5197 points)
Depth range	$0.3 \sim 5.0\,[m]$
Accuracy	$< \pm 50\,[mm]$

3D LIDAR Fig. 1(b) shows the 3D LIDAR (FX8 by Nippon Signal Co.). Table 1 summarizes the specification of the sensor we are currently using. We use two sensors to cover a 120 $[deg.]$ horizontal field of view.

ESPAR antenna Fig. 1(c) shows a product by Tsuryo Technica Corp. based on the ESPAR antenna principle. This ESPAR antenna is composed of seven monopoles, one at the center and the others on a surrounding circle, and can control its directivity by changing the load reactance electronically. Since the directivity control is very fast, the antenna is suitable as a sensor for detecting the direction of a radio transmitter. This product is designed to return one of twelve discrete directions (every 30 $[deg.]$). We use two antennas for increasing reliability with expecting the measurement of the transmitter location by triangulation.

2.2 Software configuration

The configuration of software is shown in Fig. 2. The whole software system is composed of multiple functional modules, shown by rounded rectangles in the figure. Each module is realized as an *RT component* in the *RT-middleware*

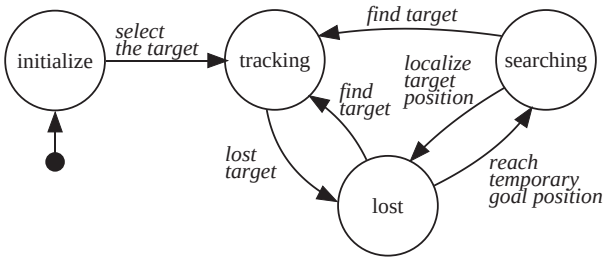


Fig. 3. State transition model.

environment [17], which supports a modularized software development. We use an implementation of RT-middleware by AIST, Japan [18]. Brief explanations of the modules are as follows.

- *Person detection*: Point cloud data is analyzed to detect person candidates using a classifier (see Sec. 3).
- *Person tracking*: Persons are tracked using a set of Kalman filters with target identification and an entropy-based filters management (see Sec. 4).
- *Target localization*: Possible target positions are determined using ESPAR antenna data in case of lost target (see Sec. 5).
- *Mapping*: A 2D probabilistic occupancy map is constructed and maintained used for path planning.
- *Path planning*: A path to follow the target is generated considering surrounding obstacles and other persons using an RRT-based path planner [19].

The first three modules have been developed in this research. The details are explained in the corresponding sections.

2.3 State transition model for robust tracking

Robustness is one of the important properties in any robotic task execution. In following a specific person, the robot has to be able to re-track the person even after losing him/her due to, for example, occlusions and out-of-view cases. Recognition of failure states and recovery from there are therefore important. Since recovery operations are usually different from those in normal situations, we introduce a state transition model.

We define the following four states (see Fig. 3):

- *Initialize*: The system starts from this state. It detects person candidates, initiates a Kalman filter for each candidate, calculates the *tracking probability* that each filter is tracking the target using ESPAR antenna data, selects the one which has the largest probability, and transits to the *tracking* state. This selected Kalman filter is called *tracking filter*.

– *Tracking*: The system is tracking a target person. That is, the tracking filter can localize the target; its positional uncertainty is within a certain range and the target probability is high enough. If one of these conditions is violated, move to the *lost* state.
– *Lost*: The system temporarily loses the target person. The robot moves to the point where the target was lost while keeping checking surrounding persons' identities. If the target is found again, move back to the *tracking* state. If the robot reaches that point without finding the target, move to the *searching* state.
– *Searching*: The system considers the target is completely lost; most probably the target is out of sight of the LIDAR or fully occluded by objects or other persons. The robot stops to calculate a target existence probabilistic grid map, which is used to limit the possible area where the target may exist. If the target is found again in this state, move directly to the *tracking* state. If not, move back to the *lost* state with an estimate of possible target position.

3 Person Detection

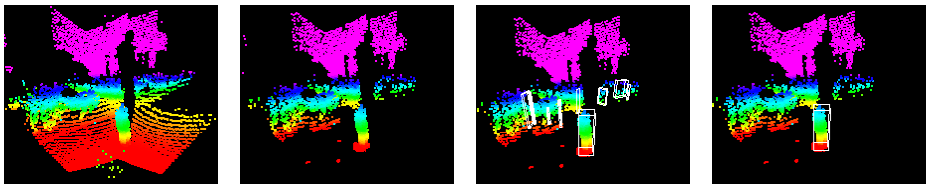
We apply a person detection method [11], originally developed for a high-definition LIDAR (i.e., Velodyne), to point cloud data obtained by our 3D LIDAR. The person detection is performed in the following steps:

1. *Segmentation*: All 3D points are mapped onto cells of a 2D grid. A cell is classified as an *object cell* if the difference between the highest and the lowest point is larger than a threshold (currently, $0.2\,[m]$). Otherwise, a cell is classified as the ground.
2. *Clustering*: A simple Euclidean clustering is applied to all object cells to form *object candidates*. A parallelepiped is then fitted to each candidate by the Rotating Calipers method [20]. If the dimensions of the parallelepiped satisfy the following conditions (unit is $[m]$) :

$$0.8 \leq \mathit{height} \leq 2.0, \quad \mathit{width} \leq 1.2, \quad \mathit{depth} \leq 1.2, \tag{1}$$

   the cluster is considered a person candidate.
3. *Classification*: Each candidate is classified into a person or non-person using RealAdaBoost [21]. The feature vector is 213-dimensional and includes both shape and reflectance features. Refer to [11] for the details of the features. For training, we manually classify detected person candidates into positive and negative for collecting training data.

Fig. 4 shows an example of person detection steps from a 3D point cloud. For a test data set with 800 range images taken from a mobile robot at about 4Hz, the recall and the precision for this test data are 90.1% and 94.6%, respectively.



(a) input point cloud (b) ground plane elimination (c) candidate clusters (c) detected person

Fig. 4. Person detection steps. Colors indicate the height of 3D point data. While boxes are superimposed on candidates (in (c)) and the detected person (in (d)).

4 Person Tracking

All detected persons are tracked independently and simultaneously using a set of Kalman filters [22]. Data association between each filter and a detected person in the current frame is done based on a position proximity and an appearance similarity. In case of occlusion, data from the ESPAR antenna is also used only for the specific person to follow. Generation and deletion of each filter is managed by an entropy-based reliability evaluation.

4.1 State variable and motion model

State vector $\boldsymbol{x}_t$ is defined as a four dimensional vector with position (X_t, Y_t) and velocity $(\dot{X}_t, \dot{Y}_t)$. We use a linear motion model represented by:

$$\begin{aligned} \boldsymbol{x}_{t+1} &= A\boldsymbol{x}_t + \boldsymbol{u}, \\ A &= \begin{bmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \end{aligned} \tag{2}$$

where ΔT is the time cycle. Deviation from the linear motion is considered by the process noise $\boldsymbol{u}$, which represents acceleration terms and whose covariance matrix is empirically determined.

4.2 Data association

People identification using reflectance values The 3D LIDAR can provide the reflectance value of each measured point. The histogram of reflectance values is used as a feature for distinguishing persons. Since we cannot know the feature in advance, we adopt an on-line boosting with feature selection [23] to learn an appearance model for each filter. When a new person is detected and a new filter for the person is initiated, an initial classifier is set up. Every time a person who matches with the filter is detected, the learning step is performed with that person as a positive sample and the others as negative samples.

The classifier has N *selectors* and each selector has M weak classifiers. At each learning, all weak classifiers and the weight α_n for the nth selector are updated. The best weak classifier $h_n^{sel}(\boldsymbol{x})$, which has the smallest error, represents the nth selector. The strong classifier is obtained by the selectors:

$$h^{strong}(\boldsymbol{x}) = \text{sign}\left(\sum_{n=1}^{N} \alpha_n \cdot h_n^{sel}(\boldsymbol{x})\right). \tag{3}$$

Data association We use two kinds of information for data association in the Kalman filtering process. One is the proximity between a detected person and a filter, evaluated by the Mahalanobis distance using a part of the covariance matrix corresponding to the positional uncertainty. The other is the appearance similarity, which is measured by the error of the classifier of the filter.

Let $d_{i,j}$ and e_j be the root-squared Mahalanobis distance and the error estimate of the strong classifier of the jth filter. We define a joint likelihood for this combination as:

$$l_{i,j} = N(d_{i,j}; 0, 1)(1 - e_j), \tag{4}$$

where $N(\cdot\,; \mu, \sigma^2)$ is the normal distribution with the mean and the variance being μ and σ^2, respectively.

Data association among detected persons and filters is performed as follows. We consider only one-to-one correspondence. First, every feasible combination of a person and a filter for which the person is verified by the filter's classifier is enumerated. The best combination with the highest joint likelihood is then chosen to fix the correspondence, and its person and filter are eliminated from the correspondence candidates. This process is repeated until no persons or no filters exist. If there is a person which does not match with any filters, a new filter is initiated for that person.

4.3 Measurement update

The measurement update step of the Kalman filter is performed by either of the following processes:

- If the corresponding person is detected, the detected position with uncertainty is used for the measurement update; the observation function $h(\boldsymbol{x})$ is defined as:

$$h(\boldsymbol{x}) = H\boldsymbol{x}, \quad H = \begin{bmatrix} 1\,0\,0\,0 \\ 0\,1\,0\,0 \end{bmatrix}. \tag{5}$$

- If the corresponding person is not detected and if the filter is for the target person, the direction of the target obtained by the ESPAR antenna is used for the measurement update. The observation function is defined as:

$$h(\boldsymbol{x}) = tan^{-1}\left(\frac{Y_t}{X_t}\right), \tag{6}$$

  where (X_t, Y_t) is the predicted position of the filter.

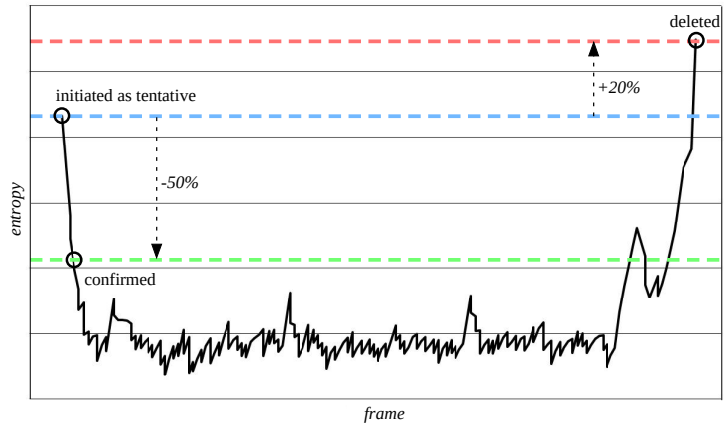


Fig. 5. An example history of entropy values and filter management.

– Otherwise, no measurement update is performed. That is, only the prediction step is performed and the uncertainty of the filter increases.

4.4 Entropy-based filters management

The management of filters, that is, generation and deletion of filters, is done by an entropy-based method [24]. They use the quadratic Rényi entropy

$$H_2(\boldsymbol{x}) = -\log \int p(\boldsymbol{x})^2 d\boldsymbol{x}. \tag{7}$$

For Gaussian distributions, we have an analytical solution given by

$$H_2(\boldsymbol{x}) = \frac{n}{2} \log 4\pi + \frac{1}{2} \log |\Sigma_t| \tag{8}$$

where Σ_t the covariance matrix and n is the dimension. In the case of the 2D tracking, Σ_t is the positional part of covariance matrix (i.e., the upper left 2×2 submatrix) of a filter and $n = 2$. This entropy is used to estimate the reliability of a filter.

Following [24], filters are considered either *tentative* or *confirmed*. When the entropy of a tentative filter drops for 50% from the initial value, it is considered to be confirmed. When the entropy of a confirmed filter becomes 20% higher than the initial value, the filter is deleted. Fig. 5 shows an example history of entropy values of a filter from its generation to deletion.

5 Target Person Identification and Localization using ESPAR Antenna

Tracking using only the range sensor may sometimes fail due to similar location and appearance between the target and others, especially in crowded cases. We therefore use ESPAR antenna data to verify if the filter for the target (called *target filter*) is actually tracking the target person. For this purpose, we associate the probability of being the target person with each Kalman filter. This probability is used for identifying the situation where the target filter loses its true target. We also use the antenna data for localizing the target in a completely lost situation.

5.1 The probability of being target

A Bayesian inference is adopted for assessing the probability that each filter is actually tracking the target person who is with the transmitter. The probability P_t^i that the ith filter is tracking the target at time t is given by

$$P_t^i = \frac{P(Z_t|A_i)P_{t-1}^i}{P(Z_t|A_i)P_{t-1}^i + P(Z_t|\overline{A_i})(1 - P_{t-1}^i)}, \tag{9}$$

$$P(Z_t|A_i) = \prod_{j=1}^{2} P(Z_t^j|A_i), \tag{10}$$

where Z_t^j is the direction of the transmitter (i.e., the target person) given by the jth ESPAR antenna at time t and A_i is the event that the ith filter is tracking the target.

The likelihood function $P(Z_t^j|A_i)$ is a Gaussian which is calculated in advance by being fitted to an actual distribution of ESPAR antenna readings. Concerning $P(Z_t^j|\overline{A_i})$, which is the probabilistic distribution of antenna readings when no target exist nearby, we model it with a uniform distribution, that is, given by $P(Z_t^j|\overline{A_i}) = 1/12$. Note that the antenna returns one of twelve directions.

The updated probabilities and the other probability P_{lost} that all filters are not tracking the target are normalized to sum up to one using the following expressions:

$$SP = \sum_i P_t^i + P_{lost}, \tag{11}$$

$$P_t^i = P_t^i/SP, \tag{12}$$

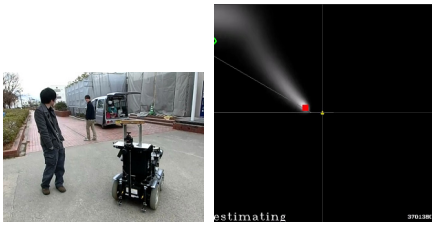
$$P_{lost} = P_{lost}/SP. \tag{13}$$

When a filter j is deleted, P_{lost} increases as follows:

$$P_{lost} = P_{lost} + P_t^j, \tag{14}$$

and decreases when a new filter is added by the probability normalization (see eq. (13)).

(a) Scene 1.



(b) Scene 2.

Fig. 6. The probabilistic target existence maps and corresponding scenes. The upward direction in the map is aligned to the heading of the robot. Red boxes in the maps indicate the highest probability cells.

5.2 Detecting lost target situation

The target is considered lost (i.e., not being tracked by the target filter) in the following two cases:

- The probability of tracking the target person by the target filter is less than a threshold (currently, 0.1).
- The entropy of the target filter is large enough to be deleted.

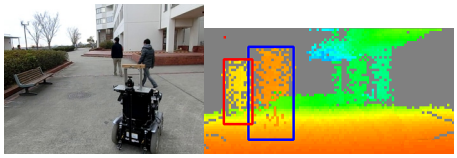
In both cases, the probabilities except that of the target filter is normalized, and if the largest probability is larger than a threshold (currently, 0.7), the corresponding filter is considered the target filter. Otherwise the system transits to the *lost* mode (see Fig. 3).

5.3 Localizing the target person using a 2D probabilistic map

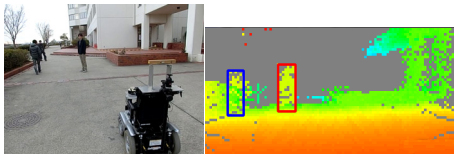
Completely lost situations usually arise when the target person is fully occluded and/or out of view of the LIDARs. The ESPAR antenna is quite suitable for finding the target in such situations. Since the direction data of the antenna include errors, especially when the robot is moving, the robot stops and accumulates antenna data for some frames using a probabilistic target existence grid map.

The probability update for each cell in the map is done in a similar way to the one used in updating the probability of a Kalman filter tracking the target (see eq. 9). We repeat this update for a certain times (currently, 10 times) and normalize the resultant grid map.

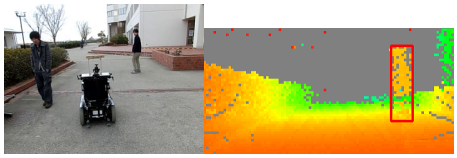
Fig. 6 shows examples of the normalized grid map and corresponding scenes. The red boxes in the figures show the highest probability cells, used as the tentative position of the target. In (a), the target location is quite accurately estimated. In (b), the target is not very well localized but its direction is correct enough for heading the robot towards the target person.



(a) A person is passing between the robot and the target.



(b) The target is correctly identified after a short occlusion.



(c) Another person is passing the left side of the robot.

Fig. 7. Short occlusion case.

6 Experimental Results

We performed specific person following experiments in our campus. This section introduces an experimental run and explains the behaviors of the system in detail.

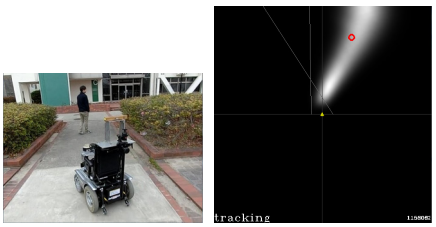
The distance traveled in this run is about 340 m and the number of total frames is 4020. The system was at *searching* mode, where the ESPAR antenna is used for searching for the target, ten times, for 271 frames in total. Person following using only LIDARs or cameras could probably had failed in one of these cases, while our system was able to continue the following behavior.

We here explain a few typical behaviors of the system. Fig. 7 shows a tracking scene with a short occlusion. All persons are detected and the target person is correctly identified. Fig. 8 shows the case where the robot loses the target out of the FOV of LIDAR and the system transits to *searching* mode. Once the target location is estimated using ESPAR antenna data, the robot moves towards there to find the target again and the system transits to *tracking* mode. Fig. 9 shows the case where the robot tracks a wrong person for a while but recovers to tracking the correct target using ESPAR antenna data.

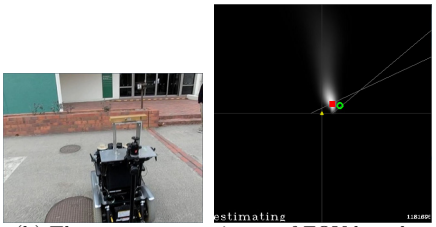
The visual sensors such as cameras and LIDARs can provide rich information for detecting people and for identifying a specific person but inherently weak to out-of-sight and misidentification situations. The ESPAR antenna can provide useful data for moving sensors and robots in the correct direction in such a situation (Figs. 8 and 9). We also tested the system at night time and it worked as well. These experimental results show the effectiveness of the proposed combination of 3D LIDAR and the ESPAR antenna.

7 Conclusions and Future Work

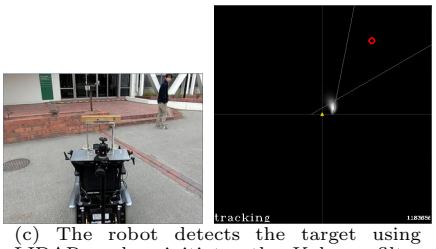
This paper has described a method of specific person detection and tracking using a 3D LIDAR and the ESPAR antenna. With the antenna that can detect



(a) The robot is following the target in the FOV of LIDAR.



(b) The target person is out of FOV but the probabilistic map (on the right) provides the target location (red box). The target filter is still running (green circle).



(c) The robot detects the target using LIDAR and re-initiates the Kalman filter tracking.

Fig. 8. Recovery from an out-of-sight situation.

the direction of a specific radio transmitter very fast, the target position can be estimated even when he/she cannot be seen in the LIDAR data, thereby making the person following behavior robust. We have developed a state transition model-based tracking method as well as robust detection and tracking procedures and implemented them on an actual mobile robot. The robot can robustly follow a specific person in difficult situations.

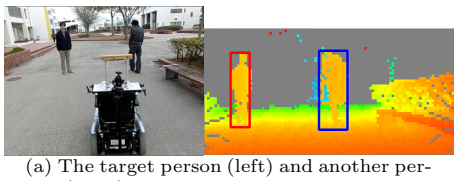
We are now planning to conduct long-run experiments in more crowded situations for quantitative evaluations. Another future work is to additionally use image data, which usually have higher resolutions, for further increasing the accuracy of detection, tracking, and identification of the target. Development of smaller antennas and transmitters is also desired for deployment.

Acknowledgment

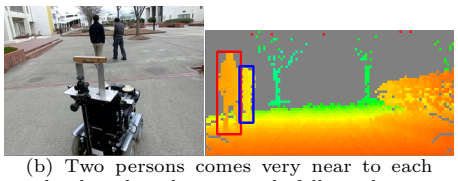
The authors would like to thank Prof. Takashi Ohira of TUT for giving advice on the ESPAR antenna. They would also like to thank the members of Active Intelligent Systems Laboratory at TUT for their supports in implementing the system. This work is supported in part by Grant-in-Aid for Scientific Research (No. 25280093) from JSPS.

References

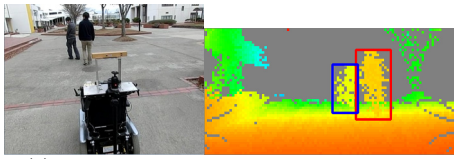
1. N. Dalal and B. Briggs. Histograms of Oriented Gradients for Human Detection. In *Proceedings of 2005 IEEE Conf. on Computer Vision and Patttern Recognition*, pp. 886–893, 2005.



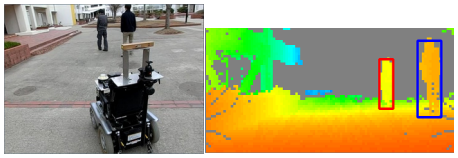
(a) The target person (left) and another person (right) walking side by side. The robot is following the target correctly.



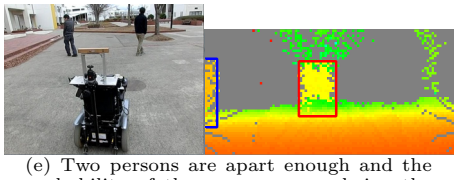
(b) Two persons comes very near to each other but the robot correctly follows the target.



(c) Two persons cross but the target is still correctly identified.



(d) The robot is temporarily tracking another person.



(e) Two persons are apart enough and the probability of the wrong person being the target decreases using ESPAR antenna data; the robot starts tracking the correct target again.

Fig. 9. Recovery from a wrong tracking.

2. J. Satake, M. Chiba, and J. Miura. Visual Person Identification using a Distance-Dependent Appearance Model for a Person Following Robot. *Int. J. of Automation and Computing*, Vol. 10, No. 5, pp. 438–446, 2013.
3. H. Tsutsui, J. Miura, and Y. Shirai. Optical Flow-Based Person Tracking using Multiple Cameras. In *Proceedings of the 2001 Int. Conf. on Multisensor Fusion and Integration for Intelligent Systems*, pp. 91–96, 2001.
4. A. Ess, B. Leibe, K. Schindler, and L.V. Gool. A Mobile Vision System for Robust Multi-Person Tracking. In *Proceedings of the 2008 IEEE Conf. on Computer Vision and Pattern Recognition*, 2008.
5. D. Schulz, W. Burgard, D. Fox, and A.B. Cremers. People Tracking with a Mobile Robot Using Sample-Based Joint Probabilistic Data Association Filters. *Int. J. of Robotics Research*, Vol. 22, No. 2, pp. 99–116, 2003.
6. K.O. Arras, O.M. Mozos, and W. Burgard. Using Boosted Features for the Detection of People in 2D Range Data. In *Proceedings of the 2007 IEEE Int. Conf. on Robotics and Automation*, pp. 3402–3407, 2007.
7. C. Premebida, O. Ludwig, and U. Nunes. Exploiting LIDAR-based Features on Pedestrian Detection in Urban Scenarios. In *Proceedings of the 12th IEEE Int. Conf. on Intelligent Transportation Systems*, pp. 18–23, 2009.
8. Z. Zainudin, S. Kodagoda, and G. Dissanayake. Torso Detection and Tracking using a 2D Laser Range Finder. In *Proceedings of Australasian Conf. on Robotics and Automation 2010*, 2010.

9. L. Spinello, K.O. Arras, R. Triebel, and R. Siegwart. A Layered Approach to People Detection in 3D Range Data. In *Proceedings of the 24th AAAI Conf. on Artificial Intelligence*, pp. 1625–1630, 2010.
10. L.E. Navarro-Serment, C. Mertz, and M. Hebert. Pedestrian Detection and Tracking using Three-Dimensional LADAR Data. *Int. J. of Robotics Research*, Vol. 29, No. 12, pp. 1516–1528, 2010.
11. K. Kidono, T. Miyasaka, A. Watanabe, T. Naito, and J. Miura. Pedestrian Recognition Using High-Definition LIDAR. In *Proceedings of 2011 IEEE Intelligent Vehicles Symp.*, pp. 405–410, 2011.
12. J. Satake, M. Chiba, and J. Miura. A SIFT-based Person Identification using a Distance-Dependent Appearance Model for a Person Following Robot. In *Proceedings of the 2012 IEEE Int. Conf. on Robotics and Biomimetics*, pp. 962–967, 2012.
13. P.N. Belhumer, J.P. Hespanha, and D.J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 711–720, 1997.
14. K. Sugiura, Y. Makihara, and Y. Yagi. Gait Identification based on Multi-view Observations using Omnidirectional Camera. In *Proceedings of 8th Asian Conf. on Computer Vision*, Vol. 1, pp. 452–461, 2007.
15. T. Germa, F. Lerasle, N. Ouadah, and V. Cadenat. Vision and RFID Data Fusion for Tracking People in Crowds by a Mobile Robot. *Computer Vision and Image Understanding*, Vol. 114, No. 6, pp. 641–651, 2010.
16. H. Kawakami and T. Ohira. Electrically Steerable Passive Array Radiator (ESPAR) Antennas. *IEEE Antennas and Propagation Magazine*, Vol. 47, No. 2, pp. 43–50, 2005.
17. N. Ando, T. Suehiro, and T. Kotoku. A Software Platform for Component Based RT System Development: OpenRTM-aist. In *Proceedings of the 1st Int. Conf. on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAR '08)*, pp. 87–98, 2008.
18. OpenRTM. `http://openrtm.org/openrtm/en/`.
19. 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, 2012.
20. G.T. Toussaint. Solving Geometric Problems with the Rotating Calipers. In *IEEE Mediterranean Electrotechnical Conf.*, pp. 1–8, 1983.
21. R.E. Schapire and Y. Singer. Improved Boosting Algorithms Using Confidence-Rated Projections. *Machine Learning*, Vol. 37, No. 3, pp. 297–336, 1999.
22. G. Welch and G. Bishop. An Introduction to the Kalman Filter. Technical Report TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill, 1995.
23. H. Grabner and H. Bischof. On-line Boosting and Vision. In *Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition*, Vol. 1, pp. 260–267, 2006.
24. S. Jurić-Kavelj, I. Marković, and I. Petrović. People Tracking with Heterogeneous Sensors using JPDAF with Entropy Based Track Management. In *Proceedings of the 5th European Conf. on Mobile Robots*, pp. 31–36, 2011.