2025 IEEE/SICE International Symposium on System Integration (SII)
January 21-24, 2025, Munich, Germany

LiDAR-based Pedestrian Tracking Adapting to Sparse Point Cloud Utilizing Interacting Multiple Model*

Masanori Imoto[1], Haziq Muhammad[2], Kazuma Sekiguchi[3], Zool Hilmi Ismail[4] and Kenichiro Nonaka[5]

***Abstract*— The vehicle navigating through narrow and crowded environments requires detailed shape information of the surrounding pedestrians for collision avoidance. While Light Detection and Ranging (LiDAR) is highly effective at measuring the position of objects, its performance diminishes as the distance between the LiDAR and the objects increases. The number of data points acquired decreases, leading to a less informative reconstruction of the object's pose and shape. To address this issue, this study proposes a pedestrian tracking method that involves constructing multiple pedestrian models and estimating the appropriate model parameters from likelihoods using the Interacting Multiple Model, based on the number of point clouds. We prepare three models: ellipse, bounding box, and point cloud center of gravity models. The ellipse and bounding box models estimate pose and size, while the point cloud center of gravity model estimates pose. The elliptical model uses Random Sample Consensus to determine model parameters that suppress arm swing and body sway during walking. Through experimental validation, this method effectively demonstrated its ability to continuously track pedestrians, including those with only a few acquired data points from a pedestrian located far from the LiDAR, while accurately estimating their pose and size.**

I. INTRODUCTION

In recent years, obstacle recognition has become essential in the research on autonomous mobility. Accurately determining the centroid position of obstacles is crucial when selecting paths through narrow environments with a mixure of obstacles [1]. Cameras and Light Detection and Ranging (LiDAR) are widely used sensors for object recognition. The camera can obtain detailed visual information. Thus, methods for tracking pedestrians in occlusion scenarios, such as environments with mixed pedestrian traffic, have been studied [2]. However, accurately estimating distances is challenging with cameras, and privacy protection, such as preserving the appearance, is essential [3]. LiDAR can accurately measure objects' position by acquiring the surrounding environment as a point cloud. Research using LiDAR includes the development of autonomous mobile robots for navigation [1]. Tracking methods based on the movements and appearances of pedestrians in crowded environments have also been studied [4], [5]. Furthermore, some research has fused LiDAR, cameras, and other sensors to improve accuracy and robustness [6].

*This work was supported by JSPS KAKENHI Grant Number 22K19801

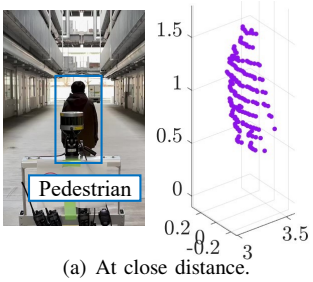
[1][2]Graduate School of Integrative Science and Engineering, Tokyo City University, Setagaya, Tokyo, Japan. {g2481008, g2291008}@tcu.ac.jp

[3][4][5]Faculty of Science and Engineering, Tokyo City University, Setagaya, Tokyo, Japan. {ksekiguc, ismail, knonaka}@tcu.ac.jp

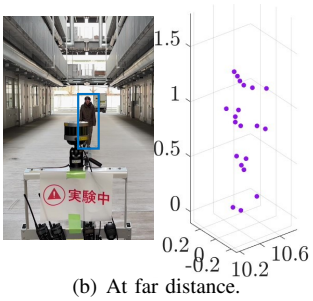
[4]Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia. ismail@tcu.ac.jp

In addition, it is necessary to estimate the shape and pose of pedestrians accurately. For shape estimation, some methods include representing the contour of a pedestrian using the variance of the pedestrian point cloud [7], tracking pedestrians using the cyclic nature of walking [8], and applying a bounding box (BB) [9], which can utilize both shape and pose information. Information about pedestrians obtained from sensors can be effectively and easily acquired by approximating them to simple shapes such as BBs. Additionally, some methods approximate the outline of pedestrians to cylinders [4]. These approximation models calculate the shape of pedestrians based on the point cloud radiated by LiDAR, making it possible to treat them as more accurate centroid positions than the point cloud centroid. However, fitting approximation models accurately to pedestrian point clouds at each time step is challenging due to arm swings and body sway. Therefore, methods have been proposed to track pedestrians using Random Sample Consensus (RANSAC) [10] to perform accurate model fitting [11], [12]. These studies suppress arm swings and body sway during walking by fitting pedestrian point clouds to models using the procedure [11]. These methods are effective for pedestrians close to LiDAR, because there is enough density of point clouds available, as demonstrated in Fig. 1(a). Collision avoidance at close range may lead to abrupt vehicle maneuvers, potentially causing discomfort to both passengers and pedestrians. If pedestrians can be detected from a distance, it becomes possible to exploit the necessary information to enable predictive avoidance. However, as shown in Fig. 1(b), it becomes difficult to capture the shape of far-off pedestrians since the point clouds become sparse. Hence, a tracking method is necessary to acquire pedestrian shapes and use appropriate models depending on the distance.

The Interacting Multiple Model (IMM) method based on the Kalman filter (KF) [14] can play an effective role as a means of accomodating to this problem. This method integrates state estimation values from multiple models based on their reliability. These models, characterized by different features, consist of multiple types, and methods using models distinguished by velocity dimensions such as constant acceleration or constant velocity have been proposed [15]–[19]. In research using LiDAR, surveillance systems such as vehicle monitoring in environments prone to occlusions and safety monitoring systems in airports have been developed [20], [21]. Additionally, it is used as an object tracking system for autonomous vehicles, where methods have been proposed to classify and track pedestrians and vehicles separately [22], and a method for tracking surrounding objects using onboard



(a) At close distance.



(b) At far distance.

Fig. 1. Difference in the number of measured pedestrian point clouds by distance. The sparsity of the point cloud changes with distance.

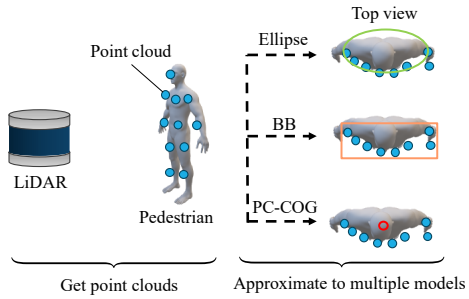


Fig. 2. LiDAR point clouds are fitted into three different models to leverage geometric features. The point cloud of the pedestrian acquired from LiDAR is model-fitted as 2D data shown in the top view.

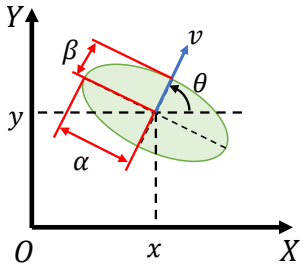


Fig. 3. Ellipse estimation model. This model consists of 6 states including pose and size.

LiDAR has also been proposed [23]. These studies are used to account for differences in the speed and angular velocity of target objects and to classify and track multiple objects based on their shapes. Thus, the appropriate use of multiple models is expected to adapt to the fluctuation of density of point clouds.

In this study, we propose a pedestrian tracking method using IMM for LiDAR-based vehicle control. By using the IMM filter with multiple models, we aim to cover the distance-dependent measurement performance of LiDAR while seamlessly tracking changes in distance. At close range, where the point cloud density is high, we use RANSAC to obtain the shape of pedestrians based on the elliptical model, enabling accurate estimation of their positions and velocities. At longer distances, where the point cloud density decreases, we estimate the position and velocity using BB or point cloud center of gravity (PC-COG) model. Through the IMM, each model interacts with others, allowing consideration of pedestrian shapes even at long distances. Compared to the typical single-model tracking, this system can be used as a wide-area tracking system without relying highly on the measurement performance of the LiDAR, and is expected to enable flexible avoidance in vehicle avoidance control. To verify the effectiveness of tracking concerning changes in pedestrian distance, we conducted validation using a tracking system incorporating the IMM filter. In scenarios involving long-distance tracking using a fixed LiDAR, we confirmed continuous tracking of pedestrians at distances of 30 m or more.

The rest of this paper is organized as follows. Section II describes the estimation model used, section III describes the tracking system as a whole, section IV focuses on the IMM within the tracking system, and section V presents the validation. The final section concludes the study.

II. ESTIMATION MODEL

This section describes the model of pedestrians targeted for estimation. Note that it is challenging to recognize the shape of pedestrians at long distances, so we propose to track them using multiple models. After acquiring pedestrian point clouds with LiDAR, we fit them to models with different shapes, as shown in Fig. 2. The models used are elliptical, BB, and PC-COG models.

A. Pedestrian model

The pedestrian pose is represented as the center of the human body torso (x,y) and the orientation angle θ in the Cartesian coordinate system. The walking velocity is v, and the size of the body is described by the major axis α and the minor axis β, as depicted in Fig. 3. The state of the model $\boldsymbol{x}$ is defined by

$$\boldsymbol{x} := \begin{bmatrix} x & y & \theta & v & \alpha & \beta \end{bmatrix}^\top \in \mathbb{R}^6. \tag{1}$$

Furthermore, since the pedestrian's intention is uncertain, we assume it as a random walk model. Its dynamics is $\dot{x} = v\cos\theta$, $\dot{y} = v\sin\theta$, and $\dot{v}, \dot{\theta}, \dot{\alpha}$ and $\dot{\beta}$ obeys Gaussian signal. Then, the state transition model at time step k is

$$\boldsymbol{x}[k] = \boldsymbol{f}(\boldsymbol{x}[k-1]) + \boldsymbol{n}[k], \tag{2}$$

where the process noise $\boldsymbol{n}$ obeys a normal distribution $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$ with $\boldsymbol{Q} \in \mathbb{R}^{6\times6}$ denoting the covariance matrix.

B. Elliptical model for estimation

We describe the ellipse model [12] used for estimation in Fig. 3. This model approximates the torso of the pedestrian point cloud in the left part of Fig. 2, as illustrated in

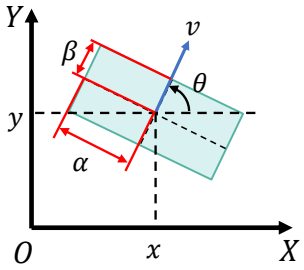


Fig. 4. Bounding box estimation model having the same variables with the ellipse model.

the upper right illustration, and estimates the state $\boldsymbol{x}_\mathrm{e} := [x\ y\ \theta\ v\ \alpha\ \beta]^\top$. We assume that x, y, θ, α and β are measured via elliptical approximation at pedestrian point cloud. The measurement vector is $\boldsymbol{y}_\mathrm{e} := [x\ y\ \theta\ \alpha\ \beta]^\top$, and the observation equation is given by (4) using the coefficient matrix $\boldsymbol{C}_\mathrm{e} \in \mathbb{R}^{5\times6}$:

$$\boldsymbol{y}_\mathrm{e}[k] = \boldsymbol{C}_\mathrm{e}\boldsymbol{x}[k] + \boldsymbol{w}_\mathrm{e}[k], \tag{3}$$

where the observation noise $\boldsymbol{w}_\mathrm{e}$ obeys a normal distribution, $\boldsymbol{w}_\mathrm{e} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_\mathrm{e})$, and $\boldsymbol{R}_\mathrm{e} \in \mathbb{R}^{5\times5}$ is the covariance matrix.

C. Bounding box model for estimation

The BB model defines the state as $\boldsymbol{x}_\mathrm{b} := [x\ y\ \theta\ \alpha\ \beta]^\top$, similar to the elliptical model in Section II-B. The BB model is shown in Fig. 4. We assume that x, y, θ, α and β are measured via BB approximation at pedestrian point cloud. Then, the measurement vector is $\boldsymbol{y}_\mathrm{b} := [x\ y\ \theta\ \alpha\ \beta]^\top$, and the observation equation is given by (4) using the coefficient matrix $\boldsymbol{C}_\mathrm{e} \in \mathbb{R}^{5\times6}$ as

$$\boldsymbol{y}_\mathrm{b}[k] = \boldsymbol{C}_\mathrm{e}\boldsymbol{x}[k] + \boldsymbol{w}_\mathrm{b}[k], \tag{4}$$

where the observation noise $\boldsymbol{w}_\mathrm{b}$ obeys a normal distribution, $\boldsymbol{w}_\mathrm{b} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_\mathrm{b})$, and $\boldsymbol{R}_\mathrm{b} \in \mathbb{R}^{5\times5}$ is the covariance matrix.

D. Point cloud center of gravity model for estimation

In PC-COG model, the state is defined as $\boldsymbol{x}_\mathrm{p} := [x\ y\ \theta\ v\ \alpha\ \beta]^\top$. We assume that the centroid position (x, y) is measured and the measurement vector is $\boldsymbol{y}_\mathrm{p} := [x\ y]^\top$, and the observation equation is given by (5) using the coefficient matrix $\boldsymbol{C}_\mathrm{p} \in \mathbb{R}^{2\times6}$:

$$\boldsymbol{y}_\mathrm{p}[k] = \boldsymbol{C}_\mathrm{p}\boldsymbol{x}[k] + \boldsymbol{w}_\mathrm{p}[k], \tag{5}$$

where the observation noise $\boldsymbol{w}_\mathrm{p}$ obeys a normal distribution, $\boldsymbol{w}_\mathrm{p} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_\mathrm{p})$ and $\boldsymbol{R}_\mathrm{p} \in \mathbb{R}^{2\times2}$ is the covariance matrix. We add an offset by the length of the minor axis of the elliptical model to bring the COG closer to the torso center. Since the ellipse size α and β are unobservable, instead of estimating them, we assumed that $\dot{\alpha}$ and $\dot{\beta}$ are Gaussian signal.

III. PEDESTRIAN TRACKING

A. System's overview

The outline of pedestrian tracking system is shown as Fig. 5. The point cloud data acquired from LiDAR is pre-processed and passed to the track manager, which applies the Hungarian method [13]. The pedestrian point cloud is approximated to each model, and the estimated values are then calculated using the IMM of the proposed method. This process is repeated each time step. The estimation method using the IMM is described in the next section.

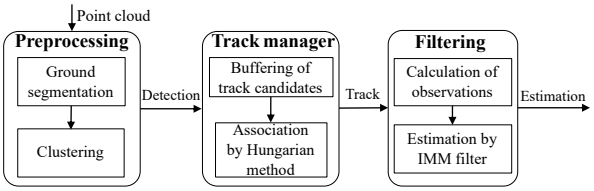


Fig. 5. Pedestrian tracking system pipeline.

B. Parameter estimation of elliptical model by RANSAC

This method follows the calculation of the elliptical model parameters as described in [12]. We treat the three-dimensional point cloud data as two-dimensional data for fitting. Let M denote the number of pedestrian point clouds, and each pedestrian point cloud is represented as (x^i_{PC}, y^i_{PC}), where $i \in \{1, \cdots, M\}$. To approximate the torso point cloud (hereafter referred to as 'inlier') to an ellipse model, the point clouds of the arms, considered as outliers, are excluded [12]. According to [24], RANSAC can determine a reasonable number of iterations with reliability considering to effectively exclude outliers. Therefore, we predefine a maximum number of interations m based on reliability C. $\boldsymbol{N}_j$ refers to the sample set described in the following steps.

1) We create a randomly selected sample set $\boldsymbol{N}_j$ from the pedestrian point clouds (x^i_{PC}, y^i_{PC}).
2) Based on optimization calculations, we compute candidate ellipse models from all sample point clouds $(x^i_s, y^i_s), i \in \boldsymbol{N}_j$. The candidate models are determined by the following binomial evaluation function:

$$J_j = \sum_{i\in\boldsymbol{N}_j} \left(s^i_j\right)^2 + \exp\left\{\sqrt{{x^i_{PC}}^2 + {y^i_{PC}}^2} - \sqrt{x_j^2 + y_j^2}\right\} \tag{6}$$

s^i_j represents the error between (x^i_{PC}, y^i_{PC}) and the ellipse model as follows, where $\tilde{x}_j = x^i_{PC} - x_j$ and $\tilde{y}_j = y^i_{PC} - y_j$:

$$\begin{aligned} s^i_j = &\left(\frac{-\tilde{x}_j \sin\theta_j + \tilde{y}_j \cos\theta_j}{\alpha_j}\right)^2 \\ &+ \left(\frac{-\tilde{x}_j \cos\theta_j - \tilde{y}_j \sin\theta_j}{\beta_j}\right)^2 - 1 \end{aligned} \tag{7}$$

The second term of (6) serves as a correction term for fitting [12], preventing the estimated center of the ellipse model from being outside the torso. We impose the following constraint, where θ' is the posture angle

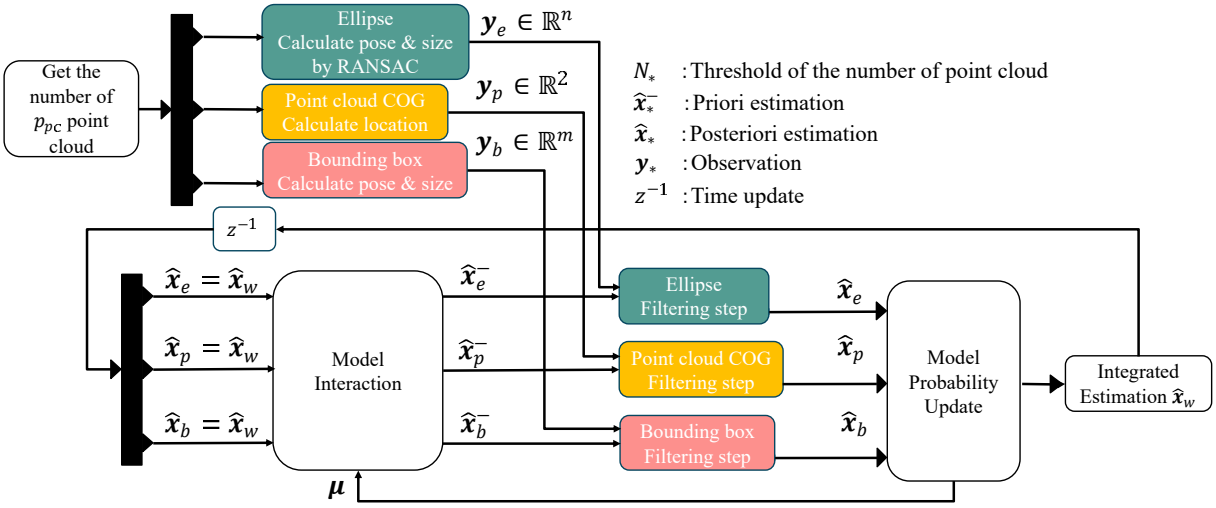


Fig. 6. Pipeline of IMM filter selecting a model for model fitting based on the number of pedestrian point clouds. The fitted model is used as the EKF observation. Then, the estimated values are integrated.

from the previous time step:

$$\alpha_j \geq \beta_j \tag{8}$$

$$\alpha_{\min} \leq \alpha_j \leq \alpha_{\max} \tag{9}$$

$$\beta_{\min} \leq \beta_j \leq \beta_{\max} \tag{10}$$

$$\theta' - \theta_{\text{th}} \leq \theta_j \leq \theta' + \theta_{\text{th}} \tag{11}$$

Here, the optimization problem for RANSAC:

$$\begin{aligned} &\text{minimize } J_j, \\ &\text{with respect to } x_j, y_j, \theta_j, \alpha_j, \beta_j, \\ &\text{subject to } (8)(9)(10)(11). \end{aligned} \tag{12}$$

3) We calculate the error evaluation value between the ellipse model obtained from step 2 and the pedestrian point clouds (x_{PC}^i, y_{PC}^i). Then, we compute the sum of evaluation values for M pedestrian point clouds using

$$W_j = \sum_{i=1}^{M} \left(s_j^i\right)^2. \tag{13}$$

We adopt the ellipse model that minimizes W_j until k iterations are completed. Points (x_{PC}^i, y_{PC}^i) exceeding the threshold d_{th} are considered outliers, indicating point clouds undesirable for fitting, such as those from the arms. The following pseudo-code is summerizes this procedure:

Algorithm 1 Estimate ellipse model by RANSAC(x_{PC}, y_{PC}, m)

for 1 to k **do**
  $(x_s^i, y_s^i) \leftarrow \text{rand}(x_{PC}, y_{PC})$, where $i \in \boldsymbol{N}_j$
  $(x_j, y_j, \theta_j, \alpha_j, \beta_j) \leftarrow$ Solve (12)
  Get evaluation value W_j using (13)
end for
$W_{\text{m}} \leftarrow \min(W_j)$
return W_{m}

IV. IMM for Elliptical, BB, and PC-COG Models

In this section, an IMM estimation method is presented where the best estimate value is calculated from multiple models based on their likelihood. Specifically, we estimate the pedestrian model using three different approaches: ellipse, BB, and PC-COG. Then, KF based method is employed for the estimation process.

A. Proposed IMM filter

The process of the IMM filter is illustrated in Fig. 6. The IMM consists of the following four steps. Here, we focus on the transitions from model i to j using the indices i, j, which indicate one of the three models. The equations for the covariance matrix $\boldsymbol{P}_*$ are omitted. For detailed equations, refer to [25].

1) Model Interaction:
The estimated state of the model $\hat{\boldsymbol{x}}_j$ is mixed based on the model probabilities $\boldsymbol{\mu} \in \mathbb{R}^3$ and the model transition probabilities $\boldsymbol{\Pi}$. The transition probabilities are defined in (14), and the integrated transition probability $\mu_{i|j}$ is given by (15).

$$\boldsymbol{\Pi} = \left[\pi_{ij}\right]_{3\times 3} \tag{14}$$

$$\mu_{i|j} = \frac{\pi_{ij}\mu_i}{\sum_{i=1}^{3} \pi_{ij}\mu_i} \tag{15}$$

The estimated state of the model j is

$$\hat{\boldsymbol{x}}_j^- = \sum_{i=1}^{3} \hat{\boldsymbol{x}}_i \mu_{i|j}. \tag{16}$$

2) Filtering:
Filtering is performed using the calculated observations $\boldsymbol{y}_*$ for each model, similar to the KF. The calculation of observations switches dimensions based on the threshold number of point clouds N_*, which will be described in section IV-B.

3) Model Probability Update:
The model probability μ_j is updated based on the likelihood Λ_j of the predicted pedestrian existence probability, as defined in (17).

$$\mu_j = \frac{\Lambda_j \sum_{i=1}^{3} \pi_{ij}\mu_i}{\sum_{i=1}^{3}(\Lambda_j \sum_{i=1}^{3} \pi_{ij}\mu_i)} \tag{17}$$

4) Integration estimation:
The estimated values of each model are integrated (weighted averaged) using $\boldsymbol{\mu}$.

B. Switching observations in a sparse point cloud environment

If a sufficient number of pedestrian point clouds are not obtained, it becomes difficult to calculate the observation model for ellipses and BBs. Therefore, the observation model is switched according to the point cloud number threshold N_e. The observation dimension number n of the ellipse is switched as follows.

$$n = \begin{cases} 5 & \text{if } p_{pc} < N_e, \\ 2 & \text{otherwise.} \end{cases} \tag{18}$$

As mentioned above, the normal observation value has a dimension of 5 for pose and size, but if it falls below N_e, only the positions x, y are observed. Similarly, the dimension number m of BBs changes according to N_b. If the number of point clouds is less than N_e, the dimension of the observation noise covariance $\boldsymbol{R}$ also becomes $\mathbb{R}^{2\times2}$.

V. EXPERIMENT

A. Experiment setup

We conducted an experiment of pedestrian tracking in an outdoor corridor shown in Fig. 7. The corridor is an ideal environment for LiDAR measurement performance, characterized by a total length of over 80 meters and minimal obstructions. A LiDAR (Velodyne VLP-16) is installed on the vehicle, and the coordinate system X, Y, Z is defined as shown in the figure. A pedestrian walks along the trajectory indicated by arrows, covering a round trip of approximately 30 m for pose estimation at near and far distances. By tracking this process with the IMM filter, we confirm validity of estimates and how the transitions of multiple models appear. The noise parameters used in EKF are shown in Table I. They are approximate values calculated in advance via preliminary experiments, which are based on empirical adjustments. PC-COG model size, α and β are set to 0.17 and 0.08 m, respectively. The other parameters were set as follows: $n(\boldsymbol{N}) = 9$, $m = 100$, $C = 0.95$, $\alpha_{\min} = 0.1\,\mathrm{m}$, $\alpha_{\max} = 0.2\,\mathrm{m}$, $\beta_{\min} = 0.05\,\mathrm{m}$, $\beta_{\max} = 0.1\,\mathrm{m}$, $\theta_{\mathrm{th}} = \pi/12\,\mathrm{rad}$, $d_{\mathrm{th}} = 0.5\,\mathrm{m}$, $N_e = 40$, $N_b = 10$.

B. Results and Discussion

Figure 8 shows the estimated trajectory of the pedestrian pose estimation. The pedestrian walks while repeating stop and go, and Fig. 8(a) is stepped from $t = 0$ to 60s. Figures 8(b) and 8(c) are also roughly consistent with the pedestrian's movement.

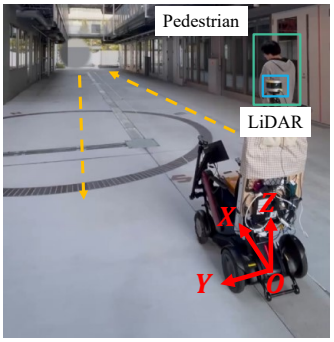


Fig. 7. Experimental scene. Pedestrian walks along the path indicated by arrows. A LiDAR located at the origin of the coordinates obtains a point cloud.

TABLE I
PARAMETER OF IMM FILTER.

Parameter	Value
$\boldsymbol{Q}_e$	diag(0,0,0.1,0.1,0.0001,0.0001)
$\boldsymbol{R}_e$	diag(9×10^{-4},9×10^{-4},$(\pi/12)^2$,9×10^{-2},9×10^{-2})
$\boldsymbol{Q}_b$	diag(0,0,0.1,0.1,0.0001,0.0001)
$\boldsymbol{R}_b$	diag(9×10^{-4},9×10^{-4},$(\pi/12)^2$,16×10^{-2},16×10^{-2})
$\boldsymbol{Q}_p$	diag(0,0,0.1,0.1,0.0001,0.0001)
$\boldsymbol{R}_p$	diag(3.6,3.6)$\times10^{-3}$
$\boldsymbol{\Pi}$	[0.99,0.005,0.005; 0.005,0.99,0.005; 0.005,0.005,0.99]

TABLE II
AVERAGE PROPORTION OF MODEL PROBABILITY BY DISTANCE

	Ellipse	**BB**	**PC-COG**
0 ∼ 10 m	**95.30**%	**0.69**%	**4.01**%
10 ∼ 20 m	**88.23**%	**10.26**%	**1.51**%
20 ∼ 30 m	**55.13**%	**41.56**%	**3.32**%
30 m ∼	**46.91**%	**31.93**%	**21.16**%

Figure 9 picks up characteristic moments such as start and goal, stop, and turning movement in time series. The probability of each model is shown in the upper left of the figure. Figures 9(b) and 9(c) are the positions where the pedestrian stopped at $X = 10, 15$ m, and the estimated values were $9.93, 14.87$m. Considering the error due to the visual observation of the pedestrian, this is a reasonable estimate. Figure 9(d) shows the turning movement, and it matches the orientation of the pedestrian in the snapshot. As shown in Figs. 9(a) to 9(f), the system was able to continuously track the pedestrian while adapting to large changes in point cloud density. Table II shows the transition of model probability with respect to distance. The proportion of elliptical models decreases as the distance increases, and other models increase. An ellipse is advantageous for tracking at close range because the shape can be accurately fitted from the point cloud distribution. In contrast, at long range, the tracking performance deteriorates because the point cloud becomes sparse and the shape becomes uncertain. Then, the model probability of BB and PC-COG increases, and the accuracy of the integrated estimation value is maintained at a high level. This is supported by the fact that the trajectories of the integrated estimation value at close and long range in Figs. 9(b)(c)(d) and Fig. 10 are biased toward each model.

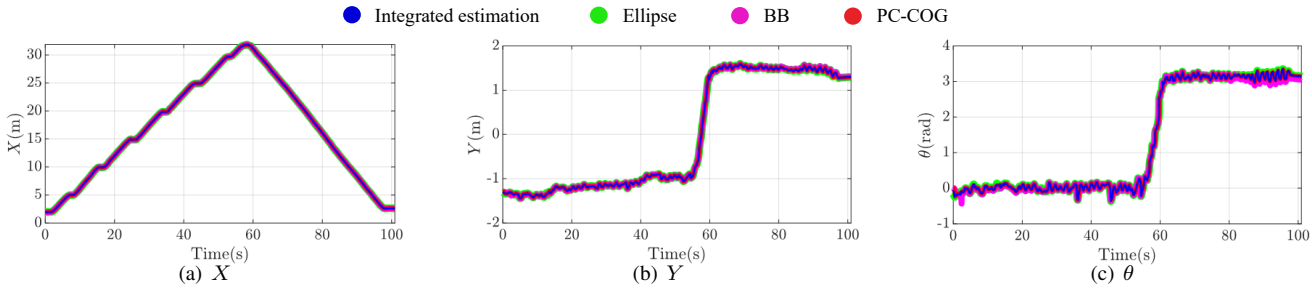


Fig. 8. Pose estimation. The observation displacements and integrated estimates for each model are shown in colors corresponding to the legend.

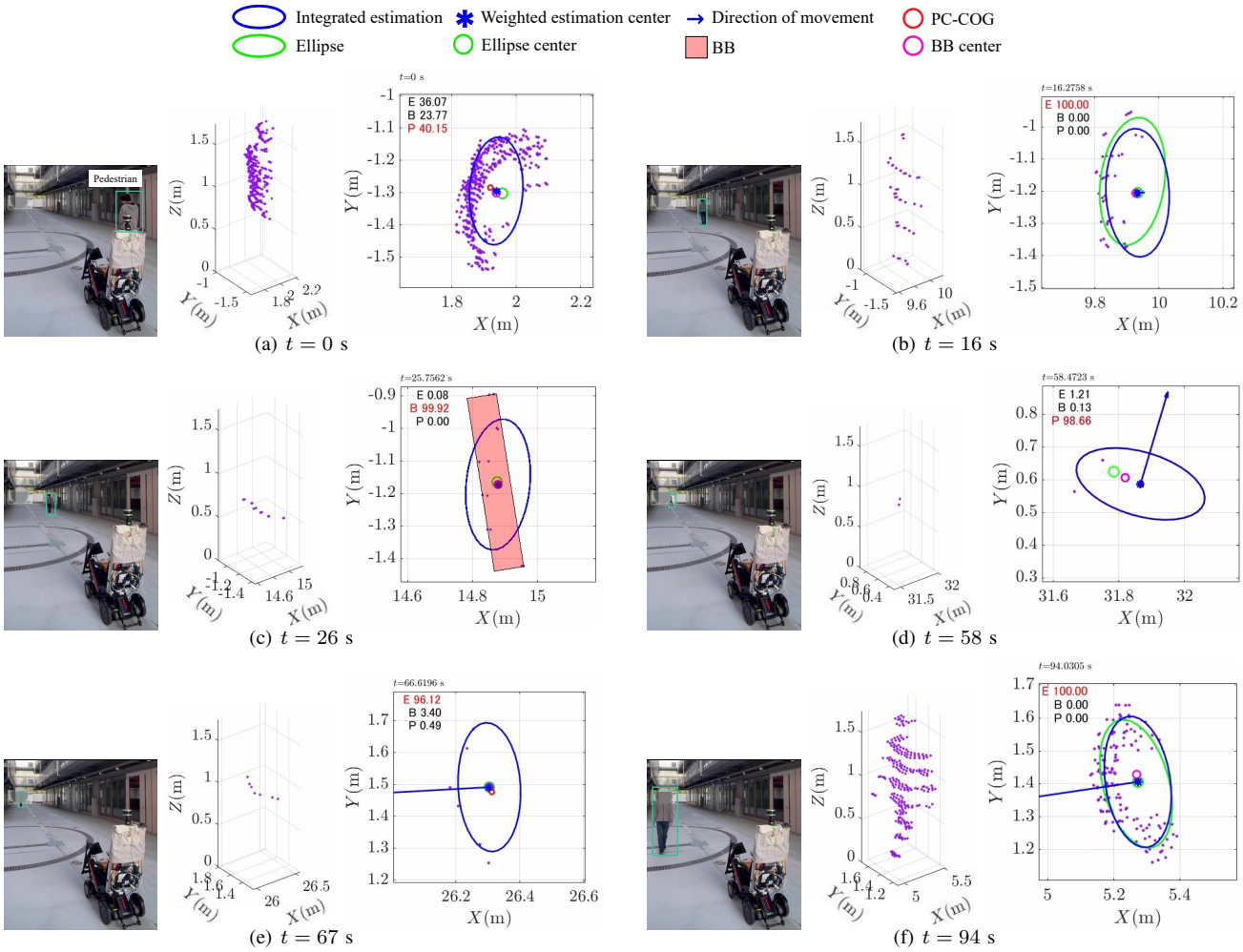
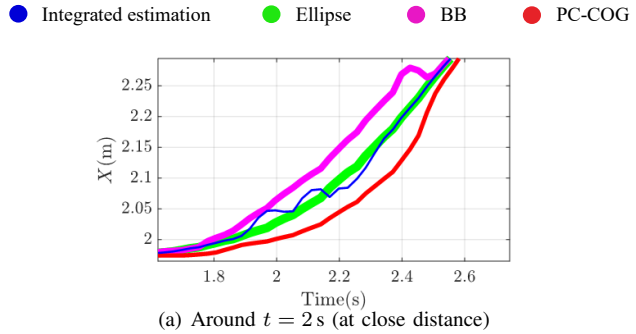


Fig. 9. Snapshot of the experiment, pedestrian point cloud data and fitting result. The model probabilities are indicated at the top left of each figure, indicating that "E" is ellipse, "B" is bouding box, and "P" is PC-COG.

For performance comparison, the tracking results with a single BB are shown in Fig. 11. With IMM, the size was maintained even when the point cloud distribution was collapsed, and the attitude angle was also correctly estimated. On the other hand, with BB, the point cloud distribution has a strong effect on the size estimation, making it difficult to estimate the correct shape.

Next, to verify the practicality when occlusion occurs, multiple pedestrians walking freely were tracked. Fig. 12 shows the situation when occlusion occurs and the tracking results of IMM. Tracks with a large number of point clouds are primarily tracked with ellipses, but PC-COG is emphasized for tracks with sparse point clouds due to occlusion. In this way, it is also adapted to partially observed point clouds. Finally, the calculation speed of the tracking system was verified through five experiments. The CPU of the verification PC was an Intel(R)Core(TM)i9-14900HX (@5.80GHz). The average calculation time was 16.9 ms,

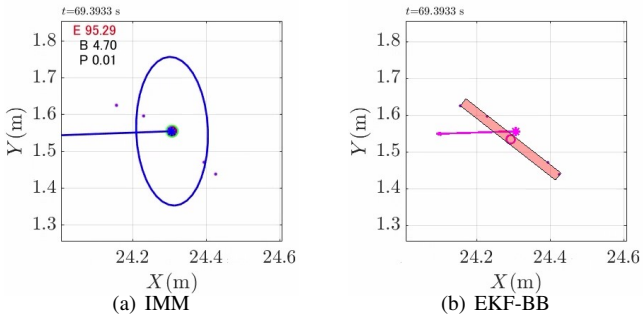
with a peak of 204.9 ms only at the moment when the number of tracked targets increased. The tracking system functions sufficiently in a steady state, but peak suppression was necessary.



(a) Around $t = 2$ s (at close distance)

(b) Around $t = 58$ s (at far distance)

Fig. 10. The trajectories of the pose estimation on the X-axis at close and far distances, as shown in the enlarged figure in Fig. 8(a). The trajectories of the integrated estimate match the trajectories of the ellipse model at close range and the PC-COG model and the BB model at far range, respectively.

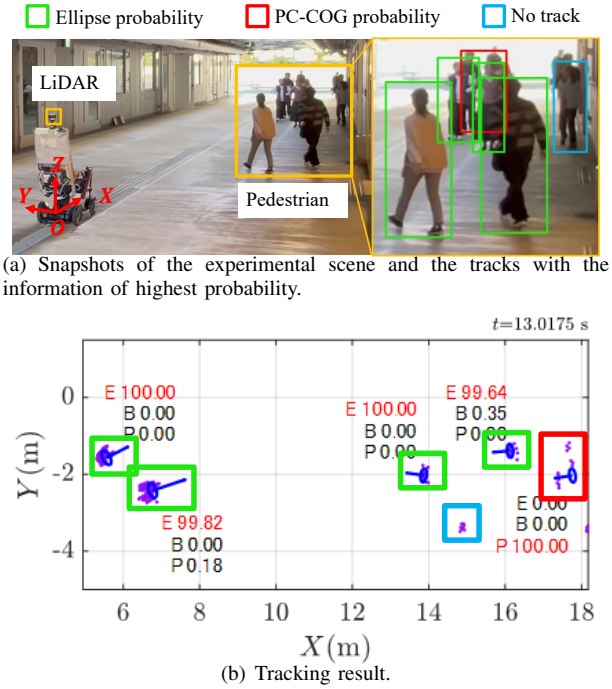


(a) IMM (b) EKF-BB

Fig. 11. Comparison of tracking in IMM and EKF-BB; pedestrian orientation and ellipse orientation are close in IMM, but not in BB, which relies on a sparse point cloud distribution, so pedestrian orientation and box orientation do not match. In fact, the model probability of the ellipse model was high, indicating that the ellipse model is superior to other models.

VI. Conclusion

In this study, we proposed a pedestrian tracking method that utilizes Interacting Multiple Model (IMM) techniques to adapt to various distances. We verified the performance of the constructed system and found that the model's probability changes based on the pedestrian's distance. At close distances, the model operates as an elliptical model, while it transitions to Bounding Box (BB) or Point Cloud-Center of Gravity (PC-POG) models as the distance increases. By



(a) Snapshots of the experimental scene and the tracks with the information of highest probability.

(b) Tracking result.

Fig. 12. Multi-pedestrian tracking result at $t = 13$ s. Pedestrians in green and red boxes are tracked, whereas others are waiting to be tracked (blue box) or neglected. Most tracked estimates rely on the ellipse model, whereas the track of the red box, which has less point cloud due to occlusion, adopts PC-COG.

integrating estimates from multiple observation models, we successfully demonstrated wide-area tracking and the system's adaptability to occlusions, where accurate pedestrian point clouds may not be available. We plan to test this method on moving vehicles and integrate it into vehicle control systems.

References

[1] Hutabarat, D., Rivai, M., Purwanto, D., and Hutomo, H., Lidar-based Obstacle Avoidance for the Autonomous Mobile Robot, 2019 12th International Conference on Information and Communication Technology and System (ICTS), pp. 197-202.

[2] Y. Lei and Z. Huang, "Research on Pedestrian Detection Algorithm Based on Monocular Vision," 2018 International Conference on Robots and Intelligent System (ICRIS), Changsha, China, 2018, pp. 161-163

[3] Ohno, M., Ukyo, R., Amano, T., Rizk, H., and Yamaguchi, H. (2024). Privacy-preserving pedestrian tracking with path image inpainting and 3D point cloud features. Pervasive and Mobile Computing, 100.

[4] Ki-In Na and Byungjae Park, Real-time 3D multi-pedestrian detection and tracking using3D LiDAR point cloud for mobile robot, vol. 45, issue. 5, 2023, pp. 836-846.

[5] L. Wang et al., "CAMO-MOT: Combined Appearance-Motion Optimization for 3D Multi-Object Tracking With Camera-LiDAR Fusion," in IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 11, pp. 11981-11996, Nov. 2023.

[6] J. Kocić, N. Jovičić and V. Drndarević, "Sensors and Sensor Fusion in Autonomous Vehicles," 2018 26th Telecommunications Forum (TELFOR), Belgrade, Serbia, 2018, pp. 420-425.

[7] Wang, W., Chang, X., Yang, J., and Xu, G. (2022). LiDAR-Based Dense Pedestrian Detection and Tracking. Applied Sciences (Switzerland), 12(4).

[8] Dolatabadi, M., Elfring, J., and van de Molengraft, R. (2020). Multiple-joint pedestrian tracking using periodic models. Sensors (Switzerland), 20(23), 1–17.

[9] Changcheng Xiao and Zhigang Luo. "Improving Multiple Pedestrian Tracking in Crowded Scenes with Hierarchical Association." Entropy, 25 (2023).
[10] Martin A. Fischler and Robert C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Communications of the ACM, vol. 24, no. 6, 1988, pp. 381-395.
[11] M. Matsuyama, K. Nonaka and K. Sekiguchi, "Estimation of pedestrian pose and velocity considering arm swing using point-cloud data," 2021 60th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), Tokyo, Japan, 2021, pp. 99-104.
[12] H. Muhammad, M. Matsuyama, K. Sekiguchi and K. Nonaka, "Pedestrian LiDAR Tracking Utilizing Elliptical Model-Based MHE Through MLESAC," 2024 European Control Conference (ECC), Stockholm, Sweden, 2024, pp. 148-153.
[13] M. L. Miller, H. S. Stone and I. J. Cox, "Optimizing Murty's ranked assignment method," in IEEE Transactions on Aerospace and Electronic Systems, vol. 33, no. 3, pp. 851-862, July 1997.
[14] H.A.P.Blom and Y.Bar-Shalom, The interacting multiple model algorithm for systems with Markovian switching coefficients, IEEE Transactions on Automatic Control, vol. 33, no. 8, 1988, pp. 780-783.
[15] Z. Jiang and D. Q. Huynh, "Multiple Pedestrian Tracking From Monocular Videos in an Interacting Multiple Model Framework," in IEEE Transactions on Image Processing, vol. 27, no. 3, pp. 1361-1375, March 2018.
[16] Z. -C. Chen and C. -W. Tang, "Robust Pedestrian Tracking Using Interactive Multiple Model Particle Filter and Feature Matching," 2018 3rd International Conference on Advanced Robotics and Mechatronics (ICARM), Singapore, Singapore, 2018, pp. 480-485.
[17] Chieh-Chih Wang, C. Thorpe and A. Suppe, "LADAR-based detection and tracking of moving objects from a ground vehicle at high speeds," IEEE IV2003 Intelligent Vehicles Symposium. Proceedings (Cat. No.03TH8683), Columbus, OH, USA, 2003, pp. 416-421.
[18] Z. Jiang and Benlian Xu, "Tracking multiple pedestrians through detection failures in videos," 2015 International Conference on Control, Automation and Information Sciences (ICCAIS), Changshu, China, 2015, pp. 420-425.
[19] Madrigal, F., and Hayet, J.-B. (2013). Evaluation of multiple motion models for multiple pedestrian visual tracking. 2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2013, 31–36.
[20] Chen, J.-F., Wang, C.-C., and Chou, C.-F. (2018). Multiple target tracking in occlusion area with interacting object models in urban environments. Robotics and Autonomous Systems, 103, 68–82.
[21] Han, B., Wang, H., Su, Z., Hao, J., Zhao, X., and Ge, P. (2023). A Gated-Recurrent-Unit-Based Interacting Multiple Model Method for Small Bird Tracking on Lidar System. Sensors, 23(18).
[22] Su, J., Chen, H., and Yang, Q. (2020). Multi-target Tracking Algorithm with Adaptive Motion Model for Autonomous Urban Driving. SAE Technical Papers, 2020-Janua.
[23] Zou, Y., Li, Y., Zhang, X., Dong, G., and Zang, Z. (2023). An Efficient Real-Time Object Detection and Tracking Framework Based on Lidar and Ins. In Lecture Notes in Electrical Engineering: Vol. 1025 LNEE.
[24] Rahul Raguram Ondrej Chum Marc Pollefeys Jiri Matas and Jan-Michael Frahm, USAC: A Universal Framework for Random Sample Consensus, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, issue. 8, 2012, pp. 2022-2038.
[25] G. Xie, L. Sun, T. Wen, X. Hei and F. Qian, "Adaptive Transition Probability Matrix-Based Parallel IMM Algorithm," in IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 51, no. 5, pp. 2980-2989, May 2021.