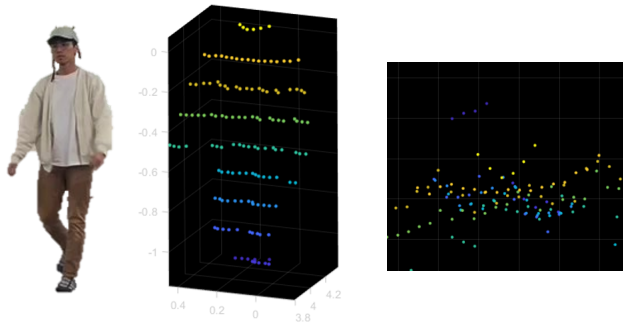
2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

Pedestrian LiDAR Tracking utilizing Elliptical Model-based MHE through MLESAC

Haziq Muhammad[1] Masato Matsuyama[2] Kazuma Sekiguchi[3] Kenichiro Nonaka[4]

***Abstract*— Pedestrian tracking using Light Detection and Ranging (LiDAR) is important to avoid pedestrians for autonomous vehicles. It is difficult, however, to measure the center position of the pedestrian directly because the point cloud data appears on the surface of the pedestrian facing the LiDAR. In addition, the arms and legs often wobble during locomotion, although the torso moves smoothly and wobbles less. Thus, in this study, to estimate the center position of the pedestrian, we approximate the torso as an ellipse model to represent the pedestrian's pose. Then, the point cloud on arms and legs that largely fluctuates is eliminated by Random Sample Consensus (RANSAC) by regarding them as outliers for the ellipse model. In addition, we propose a novel method that combines Moving Horizon Estimation (MHE) with maximum likelihood estimation sampling consensus (MLESAC) to consider the motion model of the pedestrian to prevent fitting failures. Experiments show the advantages of the proposed method.**



(a) Real (b) PC Front view (c) PC Top view

Fig. 1. Point-cloud data on pedestrian with front and top view.

I. Introduction

In recent years, the role of autonomous vehicles in environments with pedestrians, for purposes such as mobility support, has been anticipated [1], [2]. When navigating such environments, it is crucial to avoid collisions with surrounding pedestrians, making pedestrian tracking essential. Particularly in confined spaces like indoors, precise estimation of position and size is required [3]. Sensors commonly used for pedestrian tracking include Light Detection and Ranging (LiDAR) and cameras. LiDAR, as shown in Fig. 1, can perform 360-degree volumetric measurements with a single device, enabling accurate distance measurements to irradiated points compared to cameras which leads to many automated driving research using them [4]–[7]. In research on pedestrian tracking using LiDAR [8], [9], the center of mass was determined from the measured point cloud of the pedestrian's contour, estimating their position. While LiDAR can measure the distance to the irradiated points and consequently the surface of the illuminated pedestrian, directly measuring the center position is challenging. Consequently, these studies had discrepancies between the estimated center of the human body and the point cloud's center of mass, potentially leading to failures in obstacle avoidance. Moreover, techniques utilizing bounding boxes based on the measured point cloud of the tracked pedestrians were proposed [10]–[13]. While these methods consider the size of pedestrians through bounding boxes, variations in the size of the bounding box occur due to dynamically changing point clouds (PC) from arms and legs. This results in fluctuations in the center position, making accurate position estimation difficult. To address such issues, Han et al. [14] provide an example of utilizing Random Sample Consensus (RANSAC) to remove outliers, demonstrating the effectiveness of RANSAC in generating regression models for data with outliers. Similarly, in our method, PC from arms were treated as outliers and excluded by RANSAC. We performed circular approximations on the torso, which exhibits less movement during walking compared to arms and legs [15]. Subsequently, we proposed a technique that estimates the calculated circle's center and radius using a Kalman filter (KF) with these values as observations. However, discrepancies between the circle's center and the pedestrian's center position arose due to the mismatch between the circle and the torso shape. To address this, the RANSAC model was modified from a circle to an ellipse, and the torso was approximated with an ellipse for tracking [16]. In that study, the ellipse's center, orientation angle, major axis, and minor axis, computed from RANSAC, were used as observations and estimated using the Moving Horizon Estimation (MHE) method, a model-based optimization approach that accommodates observations at different time steps over the horizon. This approach significantly reduced the error between the estimated position and the pedestrian's center. Nonetheless, when approximating the point cloud using RANSAC for *each time step*, there were cases where the fit extended beyond the torso, resulting in substantial errors as illustrated in Fig. 2(a). The main problem with RANSAC itself is that it does not consider the motion model of the pedestrian. Therefore, in this study, we propose a method that employs MHE in RANSAC for 3-D LiDAR pedestrian estimation. The evaluation function of MHE can consider the pedestrian's motion model, enabling suppression of RANSAC's ellipse approximations outside the torso as shown in Fig. 2(b). The MHE evaluation function is incorporated into the optimization calculations that generate

1,2,3,4 Tokyo City University.

candidates for the RANSAC ellipse model and choose the one with the maximum likelihood. This ensures a high likelihood of RANSAC choosing the correct inliers for state estimation particularly when compared with conventional RANSAC. The effectiveness of this proposed Moving Horizon Estimation incorporated Maximum Likelihood Estimation Sampling Consensus (MHE-MLESAC) is demonstrated through real-world experiments.

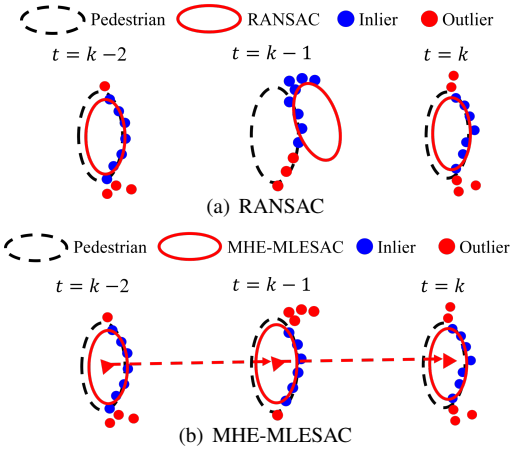


Fig. 2. PC top view comparison of conventional RANSAC and MHE-MLESAC at different times. RANSAC fails to fit the model properly with the pedestrian at $t = k - 1$ whereas, MHE-MLESAC correctly fits them because MHE-MLESAC increases the chance of proper generated models of a moving pedestrian through motion model consideration of multiple instances in horizon.

II. Estimation Target

In this study, the model considers the torso of pedestrians as an ellipse, approximating their shape, as shown in Fig. 3. The center coordinates of the torso, denoted as (x, y), correspond to the center of the ellipse. The angle θ represents the direction angle from the X-axis. Additionally, the velocities in the X and Y directions are represented as v_x and v_y, while ω represents the angular velocity, i.e. $\dot{x} = v_x$, $\dot{y} = v_y$, and $\dot{\theta} = \omega$. The major and minor axes of the ellipse are designated as α and β, respectively. For simplicity, we assume the pedestrian motion is represented by a random walk model with $\dot{v}_x = n_x$, $\dot{v}_y = n_y$, $\dot{\omega} = n_\omega$, $\dot{\alpha} = n_\alpha$, and $\dot{\beta} = n_\beta$ where n_x, n_y, n_ω, n_α, and n_β are the states' respective random noises that obeys Gaussian distribution. The state variable $\boldsymbol{x}$ is defined as:

$$\boldsymbol{x} = [x, y, \theta, v_x, v_y, \omega, \alpha, \beta]^{\mathrm{T}}. \tag{1}$$

Using Euler approximation, we can get the following discrete-time state space equation:

$$\boldsymbol{x}_k = \boldsymbol{A}\boldsymbol{x}_{k-1} + \boldsymbol{G}\boldsymbol{n}_k, \tag{2}$$

where $A \in R^{8\times 8}$ and $G \in R^{8\times 5}$ are corresponding coefficient matrices. Subscript k represents the variable at discrete time k. $\boldsymbol{n}_k := [n_x, n_y, n_\omega, n_\alpha, n_\beta]^T$ is a system noise vector that obeys Gaussian distribution, i.e. $\boldsymbol{n}_k \sim N(0, Q)$, with each velocity component being in the acceleration dimension and the components of the major and minor axes being in the velocity dimension. Additionally, $Q \in R^{5\times 5}$ represents the covariance matrix of the system noise. For the observation of the target, eventhough 3-D LiDAR PC are used, we only consider the PC's x and y positions. We calculate the deviation of individual LiDAR PC's from the ellipse model using (3):

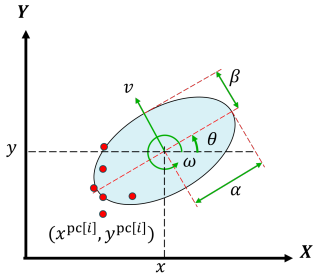


Fig. 3. Pedestrian shape model approximated by ellipse with point clouds observation. The turquoise ellipse is the pedestrian model and the red dots represent point clouds.

$$s^i = \left(\frac{\tilde{x}^i \cos\hat{\theta} + \tilde{y}^i \sin\hat{\theta}}{\hat{\alpha}}\right)^2 + \left(\frac{-\tilde{x}^i \sin\hat{\theta} + \tilde{y}^i \cos\hat{\theta}}{\hat{\beta}}\right)^2 - 1, \tag{3}$$

where superscript i represents an index, $\tilde{x}^i$ and $\tilde{y}^i$ are deviations of the i^{th} PC Cartesian position from the ellipse center where $\tilde{x}^i = x^{\mathrm{pc[i]}} - x$ and $\tilde{y}^i = y^{\mathrm{pc[i]}} - y$

III. Moving Horizon Estimation (MHE)

MHE is a model-based filter that calculates the optimal state by taking into account measurements taken in the past during a finite time period known as horizon, H. By minimizing Mahalanobis distance of the noise, we can estimate the maximum likelihood estimation of the target state. Since the optimization is numerical, constraints can also be considered to give realistic estimations based on the system's physical limitations [17], [18]. The evaluation function usually consists of three main terms which are the state error term, measurement error (innovation) term and arrival cost:

$$\begin{aligned} J(\hat{\boldsymbol{x}}_{T-H+1:T}) = &\sum_{k=T-H+1}^{T-1} ||\hat{\boldsymbol{x}}_k - \boldsymbol{A}\hat{\boldsymbol{x}}_{k-1}||^2_{(G^\dagger)^T Q^{-1} G^\dagger} \\ &+ \sum_{k=T-H+1}^{T} ||\boldsymbol{y}_k - h(\hat{\boldsymbol{x}}_k)||^2_{\boldsymbol{R}^{-1}} \\ &+ ||\hat{\boldsymbol{x}}_{T-H+1} - \boldsymbol{x}^-_{T-H+1}||^2_{(\boldsymbol{P}^-_{T-H+1})^{-1}}, \end{aligned} \tag{4}$$

where the current time is denoted as T, H represents the horizon length, the hat symbol $\hat{\boldsymbol{x}}$ represents the optimization variable and the subscript expression $T - H + 1 : T$ means from time $k = T-H+1$ until T. The first term represent the state prediction error evaluating the Mahalanobis distance for the system noise $\boldsymbol{n}_k^T Q^{-1} \boldsymbol{n}_k$. $\boldsymbol{n}_k = G^\dagger(\hat{\boldsymbol{x}}_k - \boldsymbol{A}\hat{\boldsymbol{x}}_{k-1})$ where $G^\dagger$ represents the pseudo inverse matrix, given by $G^\dagger = (G^T G)^{-1} G^T$. The second term evaluates the Mahalanobis

distance of the observation noise for the observation model $\boldsymbol{y} = h(\boldsymbol{x}) + \epsilon$ where h is a vector valued function and ϵ is the Gaussian measurement noise with covariance matrix $\boldsymbol{R}$. The third term represents the error of the optimization states $\hat{\boldsymbol{x}}_{T-H+1}$ to the *apriori* based on the past states $\boldsymbol{x}^-_{T-H+1}$ where, P^-_{T-H+1} represent the prior error covariance matrix of the KF at time $T-H+1$. In our study we run the KF simultaneously with the MHE estimator to estimate x^- and P^-. The optimization problem of MHE is to minimize J with respect to $\hat{\boldsymbol{x}}_{T-H+1:T}$.

IV. MLESAC for pedestrian PC fitting

A. Overview of model estimation with MHE-MLESAC

To estimate the pedestrian in the presence of outliers, we opted to use RANSAC as it could predict regression models from an outlier present data set by considering the highest inliers percentage [19]. However, sometimes general RANSAC may generate a wrong model as it only prioritize having the largest number of inliers. This may be reliable for static cases but not robust in dynamic scenarios. To prevent this, we incorporate MHE into RANSAC where motion model of the pedestrian is also considered when choosing inliers. We extend the RANSAC problem into a likelihood maximization problem as it synergize with the MHE estimator [20]. The Maximum Likelihood Estimation Sampling Consensus (MLESAC) will choose models with the lowest negative log-likelihood value. The idea of this method is visualized in Fig. 4. The proposed method MHE-MLESAC is similar to RANSAC where it would randomly sample among the PC within the horizon to generate candidate states. Afterwards, we will evaluate their respective negative log likelihood values where the lowest valued candidate state will be chosen. This should lead to more accurate model estimations despite the presence of outliers. The generation of candidate state and evaluation of the negative log likelihood value will be explained in Chapter IV-B and IV-C, respectively.

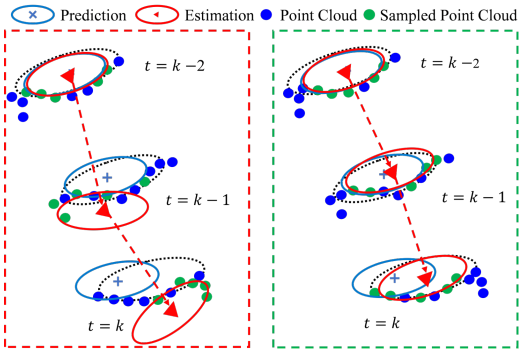


Fig. 4. MHE-MLESAC of a moving pedestrian heading downwards. The black dotted ellipse represents the real pedestrian. In the left side, poor random sampled PC subset leads to a bad estimation. However, the right side has better estimation because of good random PC samples. The estimation is better because both distance of predictions and sampled point cloud is nearer to the estimation, thus, MHE-MLESAC chooses the right side model.

B. Optimization problem for MHE-MLESAC

To establish an MHE that is capable of fitting the ellipse model to the LiDAR PC, we modified the observation term in (4) into (5).

$$\begin{aligned} &J(\hat{\boldsymbol{x}}_{T-H+1:T}; \boldsymbol{\Psi}, \boldsymbol{\mathcal{L}}) \\ &= \frac{1}{H} \sum_{k=T-H+1}^{T-1} ||\hat{\boldsymbol{x}}_{k+1} - \boldsymbol{A}\hat{\boldsymbol{x}}_k||^2_{(G^\dagger)^T Q^{-1} G^\dagger} \\ &+ \frac{1}{|\boldsymbol{\Psi}|} \sum_{k=T-H+1}^{T} R^{-1} \sum_{i \in \boldsymbol{\mathcal{L}}_k} (s^i)^2 \\ &+ ||\hat{\boldsymbol{x}}_{T-H+1} - \boldsymbol{x}^-_{T-H+1}||^2_{(\boldsymbol{P}^-_{T-H+1})^{-1}}, \end{aligned} \tag{5}$$

where the second term represents the deviation between the ellipse model and the PC $\boldsymbol{\Phi}^{\text{pc}}_k := [x^{\text{pc}}_k, y^{\text{pc}}_k]^T$ at each step. $\boldsymbol{\Psi}$ is the randomly sampled PC subset $\boldsymbol{\Psi} \subset \boldsymbol{\Phi}^{\text{pc}}_{T-H+1:T}$. $|\boldsymbol{\Psi}|$ is the total number of sampled PC, $\boldsymbol{\mathcal{L}}_k$ is an index set corresponding to the elements of $\boldsymbol{\Psi}$ extracted from $\boldsymbol{\Phi}^{\text{pc}}_{T-H+1:T}$, and $(s^i)^2$ is the square of the PC deviation defined in (3). In this case, the weight R^{-1} is a scalar. We also normalized the first term by dividing with H to avoid the term increasing disproportionately with H length. To differentiate between the major and minor axes, a constraint condition given by (6) is imposed between $\hat{\alpha}_k$ and $\hat{\beta}_k$:

$$\hat{\alpha}_k \geq \hat{\beta}_k. \tag{6}$$

Furthermore, constraints on the upper and lower limits are imposed on $\hat{\alpha}_k$, $\hat{\beta}_k$, and $\hat{\theta}_k$, as shown in (7), (8), and (9):

$$\alpha^{\min} \leq \hat{\alpha}_k \leq \alpha^{\max}, \tag{7}$$

$$\beta^{\min} \leq \hat{\beta}_k \leq \beta^{\max}, \tag{8}$$

$$\hat{\theta}^-_k - \theta^{\text{th}} \leq \hat{\theta}_k \leq \hat{\theta}^-_k + \theta^{\text{th}}, \tag{9}$$

where $\alpha^{\min}$, $\beta^{\min}$, $\alpha^{\max}$, and $\beta^{\max}$ are the minimum and maximum allowable value for $\hat{\alpha}_k$ and $\hat{\beta}_k$, respectively. They prevent from abnormal shrinking and expanding of the ellipse. $\hat{\theta}^-_k$ represents the prior estimate and θ^{th} is the allowable rotation angle. (9) is used to ensure that the ellipse's orientation angle does not change significantly. Also, to avoid the estimated ellipse center to be outside the torso, (10) is also applied:

$$\bar{r}^{\text{pc}}_k \leq \hat{r}_k, \tag{10}$$

where $\bar{r}^{\text{pc}}_k$ is the mean distance of the sampled PC subset at time k and $\hat{r}_k$ is the estimated distance of the pedestrian center to the LiDAR at time k. The optimization problem of the MHE-Ellipse is summarized as:

Optimization problem for MHE-Ellipse
minimize J in (5),
with respect to $\hat{\boldsymbol{x}}_{T-H+1:T}$,
subject to (6)(7)(8)(9)(10).

C. Evaluation of candidate states

For the MLESAC, we will generate candidate states $\hat{\boldsymbol{x}}^j_{T-H+1:T}$ with the PC subsample set $\boldsymbol{\Psi}^j$ through MHE-Ellipse where superscript j represents the index for each sample. We will determine which candidate has the highest likelihood. Firstly, we obtain the inliers subset $\boldsymbol{\Upsilon}^j$ where $\boldsymbol{\Upsilon}^j \subset \boldsymbol{\Phi}^{\mathrm{pc}}_{T-H+1:T}$. The subset $\boldsymbol{\Upsilon}^j$ contains elements $\boldsymbol{\Phi}^{\mathrm{pc}[i]}_{T-H+1:T}$ that have error $|s^i| < d^{\mathrm{th}}$ where d^{th} is the threshold value. Afterwards, we compute the negative log likelihood J using (5) where the arguments are $\hat{\boldsymbol{x}}^j_{T-H+1:T}$, $\boldsymbol{\Upsilon}^j$ and $\boldsymbol{\mathcal{M}}^j_k$ where $\boldsymbol{\mathcal{M}}^j_k$ is an index set corresponding to the elements of $\boldsymbol{\Upsilon}^j$ extracted from $\boldsymbol{\Phi}^{\mathrm{pc}}_{T-H+1:T}$. $\hat{\boldsymbol{x}}^j_{T-H+1:T}$ that produces J with the lowest value is chosen as the *aposteriori* state because it minimizes a negative log-likelihood.

V. Moving Horizon Estimation utilizing Random Sample Consensus (MHE-MLESAC)

The proposed method estimates the pedestrian ellipse model from PC using MLESAC to evaluate the likelihood through MHE. The pseudo-code is outlined in **Algorithm 1**. In the initial step when $T < H$, we estimate the model state

Algorithm 1 Ellipse model estimation based MHE-MLESAC

1: **MHE-MLESAC**($\boldsymbol{\Phi}^{\mathrm{pc}}_{T-H+1:T}$,$\boldsymbol{x}^-_{T-H+1:T}$,$j^{\max}$)
Input: Point cloud set $\boldsymbol{\Phi}^{\mathrm{pc}}$ and *apriori* state $\boldsymbol{x}^-_{T-H+1:T}$, and max trial number $j^{\max}$.
Output: Optimized state $\hat{\boldsymbol{x}}^{\mathrm{best}}_{t=T-H+1:T}$.
2: $j \leftarrow 1$
3: **while** $j \leq j^{\max}$ **do**
4: Obtain randomly sampled PC set $\boldsymbol{\Psi}^j$ from $\boldsymbol{\Phi}^{\mathrm{pc}}_{T-H+1:T}$
5: Generate a candidate model $\hat{\boldsymbol{x}}^j_{T-H+1:T}$ from $\boldsymbol{\Psi}^j$ and $\boldsymbol{x}^-_{T-H+1:T}$ by solving the **optimization problem for MHE-MLESAC**
6: Evaluate $J(\hat{\boldsymbol{x}}^j_{T-H+1:T}; \boldsymbol{\Upsilon}^j, \boldsymbol{\mathcal{M}}^j)$
7: **if** $j == 1$ **then**
8: $\hat{\boldsymbol{x}}^{\mathrm{best}}_{T-H+1:T} \leftarrow \hat{\boldsymbol{x}}^1_{T-H+1:T}$
9: $J^{\mathrm{best}} \leftarrow J^1$
10: **else**
11: **if** $J^j < J^{\mathrm{best}}$ **then**
12: $\hat{\boldsymbol{x}}^{\mathrm{best}}_{T-H+1:T} \leftarrow \hat{\boldsymbol{x}}^j_{T-H+1:T}$
13: $J^{\mathrm{best}} \leftarrow J^j$
14: **end if**
15: **end if**
16: $j \leftarrow j + 1$
17: **end while**

$\boldsymbol{x}$ through KF. When $T \geq H$ we estimate $\boldsymbol{x}$ using MHE-MLESAC. The algorithm consists of 6 main steps to estimate the state of the pedestrian. The input data would be all the pedestrian PCs $\boldsymbol{\Phi}^{\mathrm{pc}}_{T-H+1:T}$ and the *apriori* state $\hat{\boldsymbol{x}}^-_{T-H+1:T}$ within the horizon, and the maximum trial number $j^{\max}$ for MHE-MLESAC. The output of this algorithm is the new state estimation $\hat{\boldsymbol{x}}^{\mathrm{best}}_{T-H+1:T}$. The steps of the algorithm is explained below:

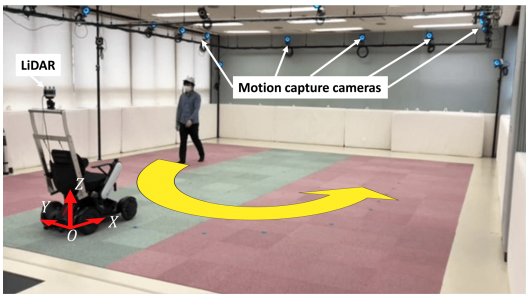


Fig. 5. The experiment environment. The LiDAR sensor is equipped on the stationary wheelchair at the origin. Motion capture cameras are equipped in the room. A pedestrian moved in the direction of yellow arrow. The 3D Cartesian coordinate system is shown in red.

1) In Line 4, randomly sample all PC $\boldsymbol{\Phi}^{\mathrm{pc}}_{T-H+1:T}$ that is within the horizon to obtain subset $\boldsymbol{\Psi}^j$.
2) In Line 5, use the sampled PC $\boldsymbol{\Psi}^j$ obtained in Line 4 and *apriori* state $\boldsymbol{x}^-_{T-H+1:T}$ as an initial prediction to determine candidate ellipse models $\hat{\boldsymbol{x}}^j_{T-H+1:T}$ through MHE-Ellipse.
3) In Line 6, evaluate the value $J(\hat{\boldsymbol{x}}^j_{T-H+1:T}; \boldsymbol{\Upsilon}^j, \boldsymbol{\mathcal{M}}^j_k)$.
4) In Lines 7 to 15, if $j = 1$, J^{best} and $\hat{\boldsymbol{x}}^{\mathrm{best}}_{T-H+1:T}$ are initialized. Later, if a better candidate model is found, J^{best} and $\hat{\boldsymbol{x}}^{\mathrm{best}}_{T-H+1:T}$ are updated with this new candidate.
5) Repeat Line 4 to 16 until $j^{\max}$ is exceeded.
6) Finally, the algorithm will output the *aposteriori* state $\hat{\boldsymbol{x}}^{\mathrm{best}}_{T-H+1:T}$.

VI. Offline Estimation Using Experimental Data

A. Experiment conditions

The proposed method was applied to PC obtained using LiDAR and was evaluated through offline estimation. The experimental setup is depicted in Fig. 5. The LiDAR was mounted on the top of an electric wheelchair, and measurements were taken while a pedestrian walked in a C-shaped path. The C-shape path was chosen as it covers the LiDAR's radiation angle of the pedestrian's shoulder and front/back side. In this experiment, PC with Z-coordinates below 0.6 m were removed as they corresponded to points from the legs. Additionally, to assess estimation accuracy, retroreflective markers were attached to the cap on the pedestrian's head, and the accurate position and orientation were obtained through motion capture (MC). The experimental parameters are written as follows: $|\boldsymbol{\Psi}| = 100$, $H = 10$, $j^{\max} = 100$, Q =diag$(0.1, 0.1, 0.08, 10^{-4}, 10^{-4})$, $R = 0.0001$, $\alpha^{\min} = 0.1\,\mathrm{m}$, $\alpha^{\max} = 0.2\,\mathrm{m}$, $\beta^{\min} = 0.05\,\mathrm{m}$, $\beta^{\max} = 0.1\,\mathrm{m}$, $\theta^{\mathrm{th}} = \pi/18\,\mathrm{rad}$ and $d^{\mathrm{th}} = 0.8$. To evaluate the performance of the present method, two methods were employed as comparison methods:

1) Estimation using KF with observations from bounding box (KF-BB) calculations [10].
2) Estimation using KF with observations from RANSAC calculated using circular model (KF-circle) [15].

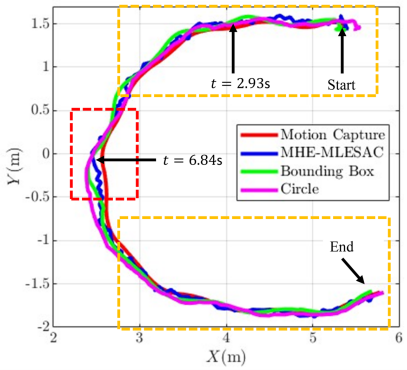


Fig. 6. Estimated trajectory of each estimated method. The yellow area represents area where the radiation of LiDAR is high because of the wider pedestrian body width facing the sensor. The red area represents area where the radiation of LiDAR is narrower because of the pedestrian orientation (shoulder side). The Pedestrian starts on the top part and ends at the bottom part as drawn by the arrows.

B. Results and discussion

Figure 6 illustrates the compared estimated positions throughout the experiment by all methods with MC on the X-Y plane. In this experiment, the pedestrian walked in a C-shaped path starting from the top part and ends at the bottom part. There are three areas boxed in Fig. 6 where PC are radiated in different sides of the body. Figure 7 is the estimation within the top orange box area at $t = 2.93$ s. The position of MHE-MLESAC and KF-BB is near to the MC whereas the shape misfit of the KF-circle method to the torso causes its centre to be far from the MC. The KF-BB is accurate in this position because the pedestrian arms are aligned with the torso. Figure 8 represents an estimation of a point within the red area at $t = 6.84$ s. Here the ellipse model was able to rotate according to the PC measurements. The correct alignment of the model makes the estimation more accurate than the other methods. Compared to Fig. 7 (c), KF-BB in Fig. 8 (c) is worse when the arm sways. Based on Fig. 9 (a) the highest positional error for all methods is around $t = 6.84$ s corresponding to the turning around $(2.54, -0.22)$. MHE-MLESAC is observed to be the least affected by the turn as the distance error remains almost the same. KF-BB distance error fluctuates strongly throughout the experiment especially during the turning in the red box area because of arm swings. Although the mean distance error of KF-BB is slightly better than MHE-MLESAC as shown in Table I, the angle error is worse. KF-circle method has the lowest accuracy because of the shape mismatch of the pedestrian's torso.

TABLE I

RMSE OF EACH ESTIMATED METHOD.

	Distance error (mm)	Abs. angular error (rad)
MHE-MLESAC	78.0	0.307
Bounding box	75.8	0.426
KF-circle	135.6	N/A

VII. CONCLUSION

In this paper, we proposed a method for estimating the pedestrian's center position by approximating the torso shape from the PC irradiated on the pedestrian and then tracking it. We utilized an elliptical model in the MLESAC fitting process, incorporating Moving Horizon Estimation (MHE) to improve the accuracy by considering the pedestrian's motion model. Experimental results demonstrated the effectiveness of the proposed method compared to other techniques.

Even though this study has the LiDAR sensor planted at the origin, with proper state transformation considering the pose of a moving vehicle, LiDAR can also be mounted to a moving vehicle for pedestrian estimation. At the current stage of this paper, this is intended to check the potential of the MHE-MLESAC tracking of the pedestrian and we are not considering the real-time potential. Future works will be to implement this algorithm in the real-time obstacle avoidance experiments.

This work was supported by JSPS KAKENHI Grant Number JP 22K19801.

REFERENCES

[1] K. Aizawa, Y. Kuramitsu, H. Matsunaga, and Y. Yasui, "Path Planning Algorithm for Micro-Mobility without Relying on Map Information," 2022 61th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), Japan, 2022, pp. 391-394.

[2] W. Maddern and P. Newman, "Real-time probabilistic fusion of sparse 3D LIDAR and dense stereo," 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Korea (South), 2016, pp.2181-2188.

[3] K. Shibata, K. Nonaka and K. Sekiguchi, "Model Predictive Obstacle Avoidance Control Suppressing Expectation of Relative Velocity against Obstacles," 2019 IEEE Conference on Control Technology and Applications (CCTA), Hong Kong, China, 2019, pp. 59-64.

[4] H. Wang and X. Zhang, "Real-time vehicle detection and tracking using 3D LiDAR," *Asian J. Control*, vol. 24, no. 3, pp. 1459-1469, May 2022.

[5] M. Omae, "Utilization of LiDAR in Autonomous Driving of Vehicles,"(in Japanese), *Journal of Measurement and Control*, vol. 59, no. 5, pp. 311-315, 2020.

[6] M. Yoshioka, N. Suganuma, K. Yoneda, M. Aldibaja, "Object Identification for Autonomous Driving Vehicles Using Omnidirectional LIDAR,"(in Japanese) in *Proceedings of the Conference on Transportation and Logistics*, vol. 2016.25, 2016.

[7] N. Akai, E. Takeuchi, T. Yamaguchi, L. Y. Morales, Y. Yoshihara, H. Okuda, T. Suzuki, Y. Ninomiya, "High-Accurate Localization INS and using Multilayer LiDAR for Autonomous Cars," *Transactions of Society of Automotive Engineers of Japan*, vol. 49, no. 3, pp. 675-681, 2018.

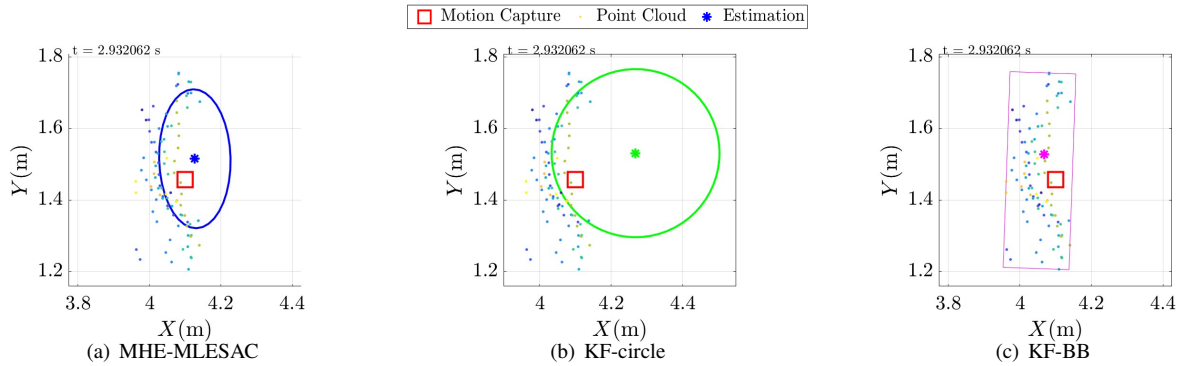
[8] J. Shackleton, B. VanVoorst and J. Hesch, "Tracking People with a 360-Degree Lidar," 2010 7th IEEE International Conference on Advanced Video and Signal Based Surveillance, Boston, MA, USA, 2010, pp. 420-426.

[9] C. Premebida et al., "A Lidar and Vision-based Approach for Pedestrian and Vehicle Detection and Tracking," in 2007 IEEE Intelligent Transportation Systems Conference, 2007.

[10] X. Zhang, W. Xu, C. Dong, and J.M. Dolan. Efficient L-Shape Fitting for Vehicle Detection Using Laser Scanners. IEEE Intelligent Vehicles Symposium, 2018.

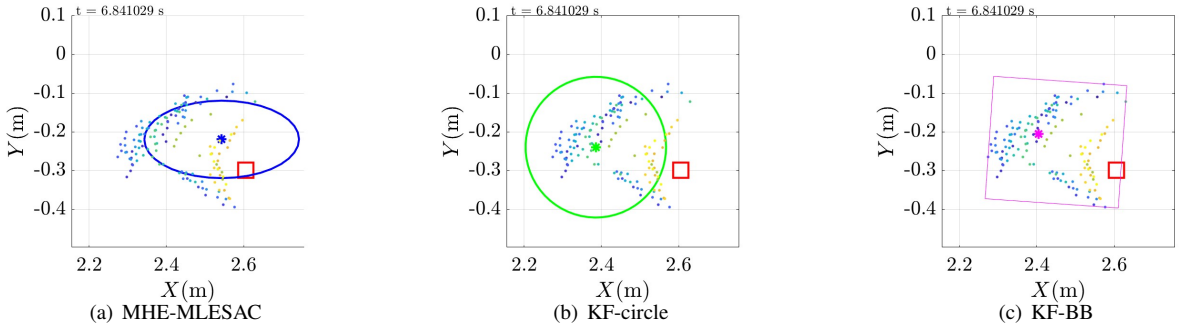
[11] Tokudome, Naruaki, et al. "Development of real-time environment recognition system using LiDAR for autonomous driving." *International Conference on ICT Robotics.*, Vol. 2017. 2017.

[12] Junxuan Zhao, Hao Xu, Hongchao Liu, Jianqing Wu, Yichen Zheng, Dayong Wu, Detection and tracking of pedestrians and vehicles using roadside LiDAR sensors, *Transportation Research Part C: Emerging Technologies*, Volume 100, 2019, pp. 68-87.



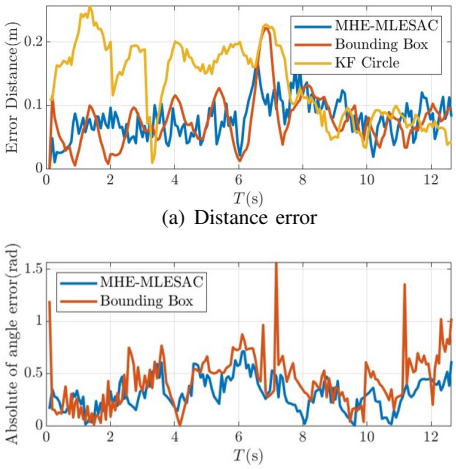
(a) MHE-MLESAC (b) KF-circle (c) KF-BB

Fig. 7. Point-cloud data and top view of fitting results at $t = 2.93\,\text{s}$. The radiation of the LiDAR laser is high because of the wider body area. MHE-MLESAC and KF-BB position estimation is near to the MC. KF-circle position is far from MC because of bad fitting of the torso.



(a) MHE-MLESAC (b) KF-circle (c) KF-BB

Fig. 8. Point-cloud data and top view of fitting results at $t = 6.84\,\text{s}$. The radiation of the LiDAR laser is lower because of the narrower body area from the sides. MHE-MLESAC is near to the MC because of proper fitting of torso and orientation estimation. KF-circle position is far from MC because of bad fitting of the torso. KF-BB accuracy deteriorates due to the swaying of arms.



(a) Distance error

(b) Angle error

Fig. 9. Positional error of each method. MHE-MLESAC maintains a small variance of distance error compared to the others. The KF-circle has the highest distance error. KF-BB is not robust to the arm sways based on the high fluctuation in (a) and the high degree of absolute angle error in (b) that .

[13] M. H¨aselich, B. J¨obgen, N. Wojke, J. Hedrich and D. Paulus, ”Confidence based pedestrian tracking in unstructured environments using 3D laser distance measurements,” 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, Chicago, IL, USA, 2014, pp. 4118-4123.

[14] M. Han, J. Kan and Y. Wang, ”Ellipsoid Fitting Using Variable Sample Consensus and Two-Ellipsoid-Bounding-Counting for Locating Lingwu Long Jujubes in a Natural Environment,” in *IEEE Access*, vol. 7, pp. 164374-164385, 2019

[15] M. Matsuyama, K. Nonaka and K. Sekiguchi, ”Estimation of pedestrian pose and velocity considering arm swing using point-cloud data,” SICE Annual Conference 2021, pp. 99-104, 2021.

[16] M. Matsuyama, K. Nonaka and K. Sekiguchi, ”Pedestrian tracking by ellipse approximation of point cloud utilizing Random Sample Consensus,” SICE Annual Conference 2022, pp. 1111-1112, 2022

[17] C. V. Rao, J. B. Rawlings, J. H. Lee, ”Constrained linear state estimation—a moving horizon approach,” *Automatica*, vol. 37, no. 10, 2001, pp. 1619-1628.

[18] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, ”Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations,” *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 246-258, Feb. 2003.

[19] M. A. Fischler and R. 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, pp. 381-395, 1981.

[20] P. H. S. Torr, A. Zisserman, ”MLESAC: A New Robust Estimator with Application to Estimating Image Geometry,” *Computer Vision and Image Understanding*, vol. 78, no. 1, 2000, pp. 138-156.