IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 8 July 2024, accepted 18 July 2024, date of publication 23 July 2024, date of current version 5 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3432590

RESEARCH ARTICLE

Human Velocity Estimation Using Kalman Filter and Least Squares With Adjustable Window Sizes for Mobile Robots

MITSUHIRO KAMEZAKI[1,2], (Member, IEEE), MICHIAKI HIRAYAMA[3], RYOSUKE KONO[3], YUSUKE TSUBURAYA[3], AND SHIGEKI SUGANO[3], (Fellow, IEEE)

[1]Department of Electrical Engineering and Information Systems, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan
[2]Future Robotics Organization, Waseda University, Shinjuku-ku, Tokyo 162-0044, Japan
[3]Department of Modern Mechanical Engineering, Waseda University, Shinjuku-ku, Tokyo 169-8555, Japan

Corresponding author: Mitsuhiro Kamezaki (kamezaki@akg.t.u-tokyo.ac.jp)

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant 19H01130 and Grant 23H01381; in part by the Waseda Research Institute for Science and Engineering, Waseda University; and in part by the Future Robotics Organization, Waseda University.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Review Committee on Research with Human Subjects of Waseda University.

ABSTRACT For autonomous mobile robots to work safely in human-coexistent environments, human-velocity estimation is essential. However, the human body periodically fluctuates to the front, rear, right, and left while walking. Also, a significant estimation error occurs due to the vibration of sensors installed in the robot. Quick trajectory adjustment requires high-accuracy and low-latency estimation, but these are in a trade-off relationship. We thus propose a human velocity estimation system (VES) using the Kalman filter (KF) and least squares (LS) with adjustable window size (AWS) to control the accuracy and latency. The VES adjusts two window sizes to calculate a system noise distribution for KF and a velocity vector for LS using a newly proposed cost function, including accuracy (direction and magnitude) and latency (time delay) costs. To select window sizes suitable for walking trajectories and individual gaits, we collected human walking data, calculated the three costs, and selected the window sizes with the minimum cost. The results of experiments using a laser range finder installed on a mobile robot indicate that the cost function could reveal window sizes to increase accuracy or reduce latency depending on walking trajectories and individual gaits, and the VES with AWS could enhance the performance of estimating human walking velocity for mobile robots.

INDEX TERMS Autonomous mobile robot, human-walking velocity estimation, Kalman filter, least squares, adjustable window size.

I. INTRODUCTION

Estimating human walking velocity with high accuracy and low latency is essential for autonomous mobile robot navigation in human-coexisting environments [1], [2], [3], [4]. Commonly used robot navigation methods, such as the velocity obstacle approach [5], [6], [7] and the potential method [8], assume that the robot can measure human velocity with high accuracy and low latency (in real-time). However, this assumption is not satisfied in real situations due to the system noise and observation noise in velocity estimation systems. Lower accuracy and higher latency prevent the robot from adequately predicting a human's path, i.e., the coordinates and timing of interference with the human, so the robot will change the path repeatedly in a short time or get stuck with selecting a non-optimal path [9].

To filter the above noises out, the Kalman filter (KF) and its derivatives are widely applied for estimating human presence and tracking humans [10], [11], [12]. KFs are effective in reducing observation noises while combining

The associate editor coordinating the review of this manuscript and approving it for publication was Pinjia Zhang.

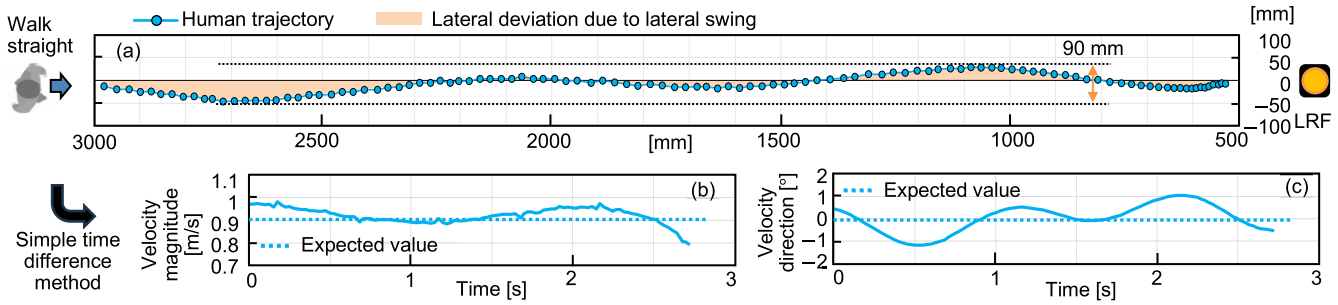

FIGURE 1. Human information estimated from LRF data when human is naturally walking straight toward LRF. (a) Human walking trajectory, which has large lateral deviation due to natural lateral swing. (b) Magnitude of walking velocity. (c) Direction of walking velocity.

other sensor data. In [13], the robot detects the legs of pedestrians using random forest based on point clouds from a laser range finder (LRF) and follows the pedestrians using the tracked trajectories derived by extended Kalman filters (EKFs). In [14], the robot follows runners at a high speed (5 m/s) in outdoor environments using KF. In [15], the robot detects and tracks a target pedestrian occluded by obstacles and other pedestrians based on KF. While many studies have focused on human tracking, there have been relatively few studies on human-velocity estimation for human-symbiotic mobile robots.

Humans have characteristics that periodically sway laterally, longitudinally, and vertically while walking due to their bio-mechanical structure [16]. Fig. 1 (a) shows human trajectory estimated using point clouds from an LRF, where we can observe a 90-mm sway in the lateral direction. Figs. 1 (b) and (c) show the velocity magnitude and direction calculated by the simple time differential of a human trajectory estimated from an LRF, where the estimated values largely deviated from the expected value (ground truth). This directly indicates that a large velocity-estimation error occurs when the velocity estimation system (VES) adopts simple derivation methods. In addition, human velocity varies with the characteristics of humans, e.g., waking gaits, environmental conditions, and vibration noises due to the robot's movement. If VES does not apply the appropriate filters to remove the above noises, the robot will suffer fatal effects, such as misunderstanding the human's intention, selecting a non-optimal path, or chattering robot behavior [17]. Filters such as KFs can mitigate the above fluctuating noises, but too strong filtering will distort the original walking trajectories, e.g., phase delay.

Some studies have developed special noise filters to reduce human fluctuating noises [13], [18]. However, they do not consider differences in individual walking gaits and various walking trajectories, nor do they control estimation accuracy and latency. Thus, developing a VES that can optimize accuracy and latency while adapting to various conditions remains challenging. As a simple but expandable system, we propose a VES with adjustable window sizes (AWS) comprising KF, least squares (LS), and cost functions. First, the VES derives the current human position using KF, which uses both LRF data and the predicted position from the estimated velocity while considering system and observation noises. The system noise is expressed as distribution, which we can adjust by changing a window size e to define the distribution. It then estimates human velocity using LS, which adopts a window size ℓ suitable for walking conditions. Conventionally, these window sizes are fixed after suitable values are found, e.g., optimizing a sliding window size for vehicle localization [19]. However, adjusting the window sizes for KF and LS affects the performance of VES: for example, quick meandering movements require real-time followability and high responsiveness, so KF needs a small noise, and LS needs a small window. On the other hand, when high anti-noise estimation is required, KF needs a large noise, and LS needs a large window.

In summary, we propose a VES that can adjust accuracy and latency. To achieve this, we estimate a human trajectory using KF and calculate a velocity using LS. A window size adjustment system adjusts e for calculating a system noise in a KF process and ℓ for calculating a velocity in an LS process, based on a cost function to learn $\{e, \ell\}$ suitable for the purpose (high accuracy, low latency, or a balance between the two). This study contributes to developing a framework for a human velocity estimation system that enables a measurement performance suitable for the target conditions and enhances the performance of mobile robot navigation.

II. RELATED AND REQUIRED WORKS

We refer to the related works on estimating walking velocity with a function to control accuracy and latency and explain the requirements for the proposed VES.

A. RELATED WORKS

The inverted pendulum model is a popular model of human-walking dynamics [20]. An inverted pendulum roughly approximates the motion of the walker's center of mass. Reference [21] revealed that lateral sway is 4.5 cm on average. Reference [22] tracked pedestrians with shoe-mounted inertial sensors based on EKF. Reference [23] recognized intentional actions of a human from the relative movements between a human and a robot using a clustering method.

Reference [24] estimated walking parameters based on a wrist-mounted inertial sensor using KF and a zero-velocity update. Reference [25] compared two walking speed estimation methods using shank- and foot-mounted inertia measurement units (IMU). Reference [26] proposed an approach to mitigate walking speed estimation errors by estimating the low-frequency noise components. Moreover, pedestrian localization and tracking systems using KF [27] and a particle filter [28] have been proposed for autonomous vehicles. Many studies have focused on human tracking, and a few have addressed human-velocity estimation. However, they used IMUs positioned on shoes or bodies, which can directly measure the acceleration of the human. In our case, we use an LRF to indirectly measure human position, which makes accurate velocity estimation extremely difficult. We thus need a VES suitable for using the human position obtained from LRF.

In information processing, stronger filtering can reduce the noise from the signal, but it will distort even the correct signal due to past time-series data, and the time-series trend will include latency. As is well known, the relationship between noise reduction and high fidelity (and realtimeness) is generally a trade-off [29]. Mobile robots require safe and efficient navigation, needing a high-accuracy and low-latency VES. At the same time, their importance or priority depends on tasks and environments, so a VES also requires a function to adjust accuracy and latency. However, to date, there have been no studies on such adjustment functions.

B. REQUIRED WORK: VES WITH AWS

As explained in Section I, to estimate a human velocity with high accuracy and low latency, our VES needs a window-size adjustment function to find window sizes suitable for walking trajectories and individual gaits. We thus developed a VES with AWS that can control the estimation accuracy and latency using our cost function. Fig. 2 shows the overview of our proposed system. It learns suitable window sizes based on costs calculated using differences from ground truth and then measures human velocity based on the selected window sizes for KF and LS.

III. VELOCITY ESTIMATION SYSTEM (VES)

We explain the theory and system configurations of our VES consisting of KF and LS, as shown in Fig. 3. Table 1 lists the symbols used in the VES. In this study, we applied a first-order Kalman filter to each axis independently.

A. OBSERVED HUMAN POSITION (SYSTEM INPUT)

Many studies on estimating human presence and tracking use RGBD image sensors installed in the environment (not robots) [12], [27], or depth sensors installed in the robot [30], [31], [32]. LRF can collect point clouds with higher accuracy, reliability, and robustness than image and depth sensors [13], [15]. In this study, the VES adopts an LRF to obtain the point clouds of pedestrians. It detects humans from point clouds and estimates their position and posture using the human detection system proposed in our previous study [3]. The human detection system first extracts 5–15 successive point clouds corresponding to a human in our experimental setting.

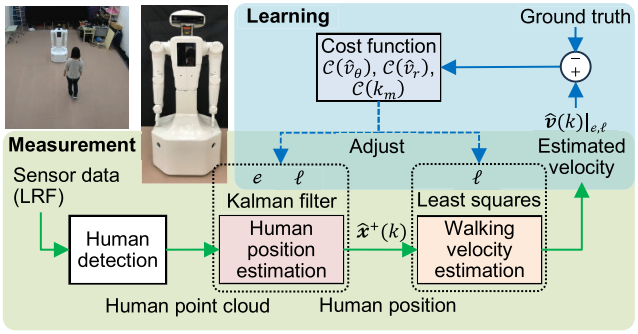

FIGURE 2. System overview of walking velocity estimation using Kalman filter and least squares with adjustable window size. The system learns suitable parameters based on costs calculated using differences from ground truth.

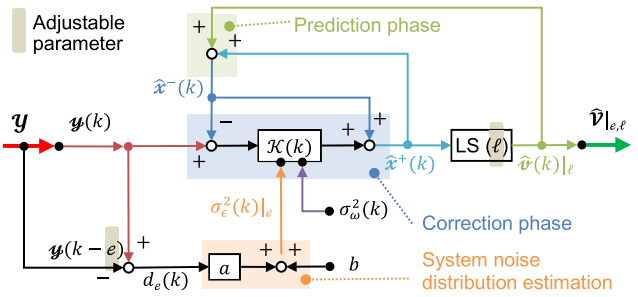

FIGURE 3. Diagram of human velocity estimation system, including KF and LS. KF consists of prediction and modification phases. a and b are constant value derived by experiments. LS means least squares.

TABLE 1. Symbols for human velocity estimation system.

k, Δt	Step $k \in \{1, \ldots, K\}$, time between each step
$\boldsymbol{y}$	Observed human trajectory
$\boldsymbol{y}(k)$	Observed position vector $\{y_1(k), y_2(k)\}^T$
$\boldsymbol{x}(k)$	State vector of human position $\{x_1(k), x_2(k)\}^T$
$\hat{\boldsymbol{x}}^-(k)$, $\hat{\boldsymbol{x}}^+(k)$	Priori and posteriori estimated position vector
ℓ	Window size for least squares ($\in \mathbb{Z}$ (integer))
e	Window size for Kalman filter ($\in \mathbb{Z}$ (integer))
$\hat{\boldsymbol{v}}(k)\|_\ell$	Estimated velocity when using ℓ $\{\hat{v}_\theta(k), \hat{v}_r(k)\}_\ell^T$
$\hat{\boldsymbol{v}}(k)\|_{e,\ell}$	Estimated velocity when using e, ℓ
$\hat{\boldsymbol{V}}\|_\ell$	Velocity of whole trajectory when using ℓ
$\hat{\boldsymbol{V}}\|_{e,\ell}$	Velocity of whole trajectory when using e, ℓ
$\epsilon(k)$, $\epsilon(k)\|_e$	System noise, system noise when using e
$\boldsymbol{\omega}(\boldsymbol{k})$	Observation noise $\{\omega_{y_1}(k), \omega_{y_2}(k)\}^T$
$\sigma_{\hat{x}^-}^2(k)$, $\sigma_{\hat{x}^+}^2(k)$	Priori and posteriori error distribution
$\sigma_\epsilon^2(k)\|_e$	System noise distribution when using e
$\sigma_\omega^2(k)$	Observation noise distribution (= 1.0)
$\mathcal{K}(k)$	Kalman gain: $0 < \mathcal{K}(k) \leq 1$
$d_e(k)$	Displacement of human using e
k_m	Time delay
$v_{\theta T}(k)$, $v_{rT}(k)$, k_{mT}	True value of direction, magnitude, & time delay
θ_m, s_m	Conditions (angle and time) of turning end
a, b	Coefficients for displacement system noise model
$\mathcal{C}$	Cost function

Then, it estimates eclipses from the extracted point clouds and calculates the human posture from the angle of long and short axes. We denote the observed human position in the target whole trajectory $\mathcal{Y} \in (\text{step } k = 1, \cdots, K)$. $\boldsymbol{y}(k)$ is the human position at step k and consists of $y_1(k)$ (longitudinal axis) and $y_2(k)$ (lateral axis), which are given by

$$\mathcal{Y} = \{\boldsymbol{y}(1), \ldots, \boldsymbol{y}(k), \ldots, \boldsymbol{y}(K)\},$$
$$\boldsymbol{y}(k) = \{y_1(k), y_2(k)\}^T. \quad (1)$$

B. STATE AND OBSERVATION EQUATIONS FOR KF

We denote the state vector of the human position $\boldsymbol{x}(k)$ as

$$\boldsymbol{x}(k) = \{x_1(k), x_2(k)\}^T. \quad (2)$$

The VES estimates the velocity using LS, which is a commonly used approximation method. The estimated velocity using the window size ℓ, $\hat{\boldsymbol{v}}(k)|_\ell$, is simply given by

$$\hat{\boldsymbol{v}}(k)|_\ell = \mathrm{LS}(\boldsymbol{x}(k-\ell), \ldots, \boldsymbol{x}(k-1), \boldsymbol{x}(k)), \quad (3)$$

where $\mathrm{LS}(\cdot)$ is the least squares function, and its argument is a time series data. To save space, please refer to [33] for a detailed calculation. The state equation is thus given by

$$\boldsymbol{x}(k) = \boldsymbol{x}(k-1) + \hat{\boldsymbol{v}}(k-1)|_\ell \cdot \Delta t + \boldsymbol{\epsilon}(k), \quad (4)$$

where $\boldsymbol{\epsilon}(k)$ is the system noise at step k, and Δt is the time gap between each step.

Then, we define the observation equation, which associates $\boldsymbol{y}(k)$ with $\boldsymbol{x}(k)$, as following:

$$\boldsymbol{y}(k) = \boldsymbol{x}(k) + \boldsymbol{\omega}(k), \quad (5)$$

where $\boldsymbol{\omega}(k)$ is the observation noise at step k. The measurement noise of a sensor is often applied to $\boldsymbol{\omega}(k)$.

C. PREDICTION AND CORRECTION PHASES FOR KF

We first explain the prediction phase. We denote the priori and posterior state vectors (i.e., human position) at step k as $\hat{\boldsymbol{x}}^-(k)$ and $\hat{\boldsymbol{x}}^+(k)$, respectively, and the priori and posterior error distributions at step k as $\sigma^2_{\hat{\boldsymbol{x}}^-}(k)$ and $\sigma^2_{\hat{\boldsymbol{x}}^+}(k)$, respectively. $\sigma^2_\epsilon(k)$ is the system noise distribution at step k. In the prediction phase, the priori state vector $\hat{\boldsymbol{x}}^-(k)$ and the prior error distribution $\sigma^2_{\hat{\boldsymbol{x}}^-}(k)$ can be given by

$$\hat{\boldsymbol{x}}^-(k) = \hat{\boldsymbol{x}}^+(k-1) + \hat{\boldsymbol{v}}(k-1)|_\ell \cdot \Delta t, \quad (6)$$
$$\sigma^2_{\hat{\boldsymbol{x}}^-}(k) = \sigma^2_{\hat{\boldsymbol{x}}^+}(k-1) + \sigma^2_\epsilon(k). \quad (7)$$

Next, for the correction phase, we denote the Kalman gain as $\mathcal{K}(k)$. As Fig. 4 shows, the posterior state vector $\hat{\boldsymbol{x}}^+(k)$ and the posterior error distribution $\sigma^2_{\hat{\boldsymbol{x}}^+}(k)$ are given by

$$\hat{\boldsymbol{x}}^+(k) = \hat{\boldsymbol{x}}^-(k) + \mathcal{K}(k)\left(\boldsymbol{y}(k) - \hat{\boldsymbol{x}}^-(k)\right), \quad (8)$$
$$\sigma^2_{\hat{\boldsymbol{x}}^+}(k) = \{1 - \mathcal{K}(k)\}\sigma^2_{\hat{\boldsymbol{x}}^-}(k). \quad (9)$$

Then, $\mathcal{K}(k)$ is updated by using

$$\mathcal{K}(k) = \frac{1}{1 + \sigma^2_\omega(k) \Big/ \sigma^2_{\hat{\boldsymbol{x}}^-}(k)}, \quad (10)$$

where $\sigma^2_\omega(k)$ is the observation noise distribution at step k.

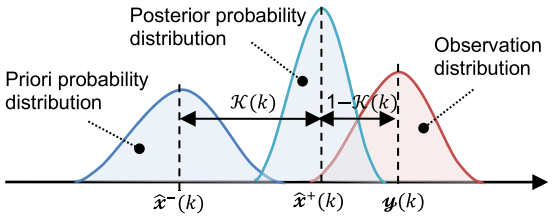

FIGURE 4. Relationships among Kalman gain $K(k)$, posterior probability, prior probability, and observation distribution.

D. VELOCITY ESTIMATION USING LEAST SQUARES

As explained earlier, $\hat{\boldsymbol{v}}(k)|_\ell$ is the estimated velocity using the window size ℓ at step k. It consists of $\hat{\boldsymbol{v}}_\theta(k)$ (direction, angle) and $\hat{\boldsymbol{v}}_r(k)$ (magnitude, speed), and is denoted by

$$\hat{\boldsymbol{v}}(k)|_\ell = \{\hat{\boldsymbol{v}}_\theta(k)|_\ell, \hat{\boldsymbol{v}}_r(k)|_\ell\}^T. \quad (11)$$

A large error would occur if we calculate the velocity using the simple time differential of the human's position. We thus adopt $\hat{\boldsymbol{x}}^+(k)$ to the LS-based velocity estimation (3). The estimated velocity $\hat{\boldsymbol{v}}(k)|_\ell$ is redefined by

$$\hat{\boldsymbol{v}}(k)|_\ell = \mathrm{LS}\left(\hat{\boldsymbol{x}}^+(k-\ell), \ldots, \hat{\boldsymbol{x}}^+(k-1), \hat{\boldsymbol{x}}^+(k)\right). \quad (12)$$

The estimated velocity in the whole trajectory $\hat{\mathcal{V}}|_l$ is finally given by

$$\hat{\mathcal{V}}|_\ell = \{\hat{\boldsymbol{v}}(1)|_\ell, \ldots, \hat{\boldsymbol{v}}(k)|_\ell, \ldots, \hat{\boldsymbol{v}}(K)|_\ell\}. \quad (13)$$

IV. ANALYSIS OF ADJUSTABLE WINDOW SIZE

Estimation accuracy and latency of the VES with KF and LS have a trade-off relationship corresponding to window sizes. To effectively reduce error and latency, we here analyze the significance of adjusting window sizes.

A. ADJUSTABLE WINDOW SIZE

From the equations in the previous subsections, we can identify the system noise distribution $\sigma^2_\epsilon(k)$, the observation noise distribution $\sigma^2_\omega(k)$, and the window size of LS ℓ as the variable parameters. We assign the measurement error of LRF to $\sigma^2_\omega(k)$, as explained before, so we develop a system to adjust $\sigma^2_\epsilon(k)$ and ℓ to control the accuracy and latency.

B. DISPLACEMENT AND SYSTEM NOISE RELATIONSHIP

We first define the displacement of human $\boldsymbol{y}(k)$ calculated using the window size e as $d_e(k)$, which is given by

$$d_e(k) = |\boldsymbol{y}(k) - \boldsymbol{y}(k-e)|. \quad (14)$$

Here, we analyze the relationship between $d_e(k)$ and the system noise distribution $\sigma^2_\epsilon(k)$. From (8), $\hat{\boldsymbol{x}}^+(k)$ is determined by $\mathcal{K}(k)$, which is the interior division ratio of $\hat{\boldsymbol{x}}^-(k)$ and $\boldsymbol{y}(k)$. $\sigma^2_\omega(k)$ is fixed to the measurement error of LRF, so $\mathcal{K}(k)$ decreases when $\sigma^2_\epsilon(k)$ is smaller while $\mathcal{K}(k)$ increases when $\sigma^2_\epsilon(k)$ is larger, as indicated in (7) and (10).

We consider a situation in which a human walks in the y_1 direction (Fig. 5 (a)). In this case, y_1 displacement $d_e(k)$ will

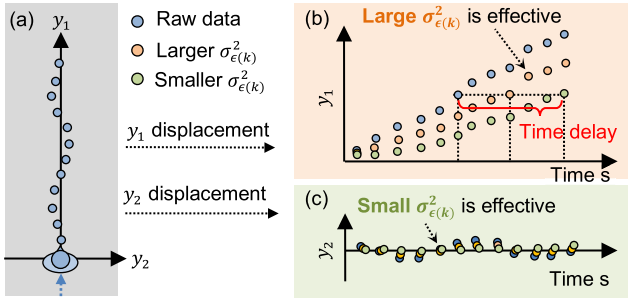

FIGURE 5. Relationship between pedestrian's displacement in y_1–y_2 coordinates and $\sigma^2_{\epsilon(k)}$. Larger $\sigma^2_{\epsilon(k)}$ is effective in responsiveness for longitudinal (y_1) axis while smaller $\sigma^2_{\epsilon(k)}$ is effective in accuracy for lateral (y_2) axis.

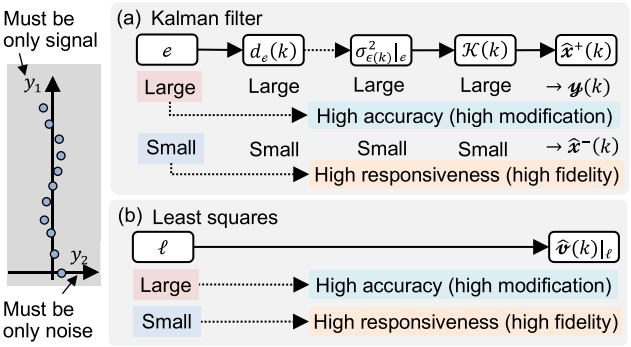

FIGURE 6. Policy of adjusting window sizes for (a) Kalman filter and (b) least squares for controlling accuracy and latency.

be large and must be only a signal to be measured (Fig. 5 (b)), so responsiveness needs to be increased, i.e., $\sigma^2_\epsilon(k)|_e$ needs to be large to trust more in $\mathcal{Y}(k)$. At the same time, the y_2 displacement $d_e(k)$ will be small and should be ideally zero since it is caused by lateral sway or sensor noises (Fig. 5 (c)), so noise suppression needs to be stronger, i.e., $\sigma^2_\epsilon(k)|_e$ needs to be small to trust more in $\hat{\boldsymbol{x}}^-(k)$.

Based on the above analysis, we summarize the effects of AWS in controlling estimation accuracy and latency by focusing on the relationship between $d_e(k)$ and $\sigma^2_\epsilon(k)|_e$. Larger $d_e(k)$ implies that the walking velocity is fast and/or the walking behavior change is large, so the system requires higher responsiveness. Alternatively, small $d_e(k)$ implies that the effect of noise, such as lateral sway on $d_e(k)$, becomes relatively significant, so the system requires higher noise suppression. Consequently, as shown in Fig. 6 (a), when $d_e(k)$ is large, the system increases the responsiveness by making $\sigma^2_\epsilon(k)|_e$ larger (trust in $\mathcal{Y}(k)$). When $d_e(k)$ is small, the system increases estimation accuracy by making $\sigma^2_\epsilon(k)|_e$ smaller (trust in $\hat{\boldsymbol{x}}^-(k)$).

C. WINDOW SIZE FOR LEAST SQUARES

We here explain how to adjust ℓ. As Fig. 6 (b) shows, a larger ℓ will provide more filtering effect, so the result has higher noise suppression, while a smaller ℓ will provide less filtering effect in real time, so the result has higher responsiveness. Thus, controlling the balance between accuracy and latency using ℓ is easier than that using e.

D. MODIFICATION AND OPTIMIZATION

Based on the above analyses, we modify the prediction phase (7) by adopting the system noise distribution when using e $\sigma^2_\epsilon(k)|_e$ as follows

$$\sigma^2_{\hat{\boldsymbol{x}}^-}(k) = \sigma^2_{\hat{\boldsymbol{x}}^+}(k-1) + \sigma^2_\epsilon(k)|_e. \tag{15}$$

Then, we modify the definition of human velocity (11) and (13) by adopting the window size for KF e as follows

$$\begin{aligned}\hat{\boldsymbol{v}}(k)|_{e,\ell} &= \left\{\hat{\boldsymbol{v}}_\theta(k)|_{e,\ell}, \hat{\boldsymbol{v}}_r(k)|_{e,\ell}\right\}^T, \\ \hat{\mathcal{V}}|_{e,\ell} &= \left\{\hat{\boldsymbol{v}}(1)|_{e,\ell}, \ldots, \hat{\boldsymbol{v}}(k)|_{e,\ell}, \ldots, \hat{\boldsymbol{v}}(K)|_{e,\ell}\right\}. \end{aligned} \tag{16}$$

The challenge here is how to find the optimal set $\{e, \ell\}$ for the VES. We explain the adjustment system in Section V.

V. WINDOW SIZE ADJUSTMENT SYSTEM

We found that the accuracy and latency can be controlled by adjusting e and ℓ. We here explain a window size adjustment system we developed.

A. FLOW OF SELECTING OPTIMAL WINDOW SIZES

As discussed in Section IV, the system uses the human trajectory $\mathcal{Y}$ and the window size $\{e, \ell\}$ and then outputs $\hat{\mathcal{V}}|_{e,\ell}$, as shown in Figs. 2 and 3. The system learns the characteristics of a human's movement in the target environments and selects $\{e, \ell\}$ to meet the required accuracy and latency. To this end, we first propose a cost function to evaluate the accuracy and latency of $\hat{\mathcal{V}}|_{e,\ell}$. Then, we collect actual human walking data, calculate the costs, and select the window size set $\{e, \ell\}$ suited to the purpose, as shown in Fig. 2.

B. COST FUNCTION

We propose a cost function for three elements to evaluate the velocity direction $\mathcal{C}(\hat{\boldsymbol{v}}_\theta)$, the velocity magnitude $\mathcal{C}(\hat{\boldsymbol{v}}_r)$, and the time delay for a human to change the direction $\mathcal{C}(k_m)$. k_m is the time delay [s]. Each element varies in magnitude and dimension, so we standardize these three values. We first calculate the root mean square error (RMSE) of the training data (e.g., trajectories of human A) for each element. We then standardize the RMSE data by $e \times \ell$ patterns of the RMSE. They are given by

$$\mathcal{C}(\hat{\boldsymbol{v}}_\theta) = \sqrt{\frac{\frac{1}{K \cdot H}\sum_H \sum_{k=1}^K \left(\hat{\boldsymbol{v}}_\theta(k) - v_{\theta T}(k)\right)^2}{\frac{1}{e \cdot \ell \cdot K \cdot H}\sum_{e,\ell}\sum_H \sum_{k=1}^K \left(\hat{\boldsymbol{v}}_\theta(k) - v_{\theta T}(k)\right)^2}}, \tag{17}$$

$$\mathcal{C}(\hat{\boldsymbol{v}}_r) = \sqrt{\frac{\frac{1}{K \cdot H}\sum_H \sum_{k=1}^K \left(\hat{\boldsymbol{v}}_r(k) - v_{rT}(k)\right)^2}{\frac{1}{e \cdot \ell \cdot K \cdot H}\sum_{e,\ell}\sum_H \sum_{k=1}^K \left(\hat{\boldsymbol{v}}_r(k) - v_{rT}(k)\right)^2}}, \tag{18}$$

$$\mathcal{C}(k_m) = \sqrt{\frac{\frac{1}{M \cdot H}\sum_H \sum_{m=1}^M (k_m - k_{mT})^2}{\frac{1}{e \cdot \ell \cdot M \cdot H}\sum_{e,\ell}\sum_H \sum_{m=1}^M (k_m - k_{mT})^2}}. \tag{19}$$

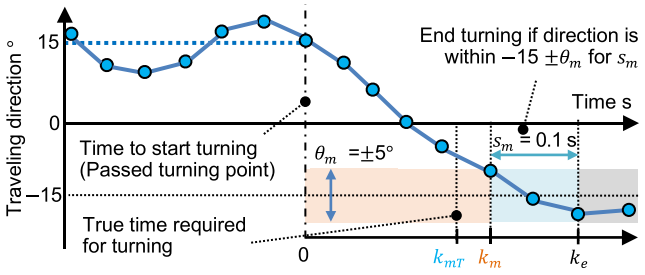

FIGURE 7. Example of determining response time for a human to change moving direction.

We denote $v_{\theta T}(k)$ as the true value of velocity direction, $v_{rT}(k)$ as the true value of velocity magnitude in the traveling direction, and k_{mT} as the true value of the time delay for a human to change the direction. K is the number of steps (data length), H is the number of training sets (participants) (= 11), and M is the number of turning patterns (= 1 or 2).

When selecting a window size $\{e, \ell\}$, we need to combine the cost of all three elements. We define the total cost $\mathcal{C}$ by multiplying the weight coefficients $\{\alpha, \beta, \gamma\}$ as

$$\mathcal{C} = \alpha \cdot \mathcal{C}\left(\hat{\mathbf{v}}_\theta\right) + \beta \cdot \mathcal{C}\left(\hat{\mathbf{v}}_r\right) + \gamma \cdot \mathcal{C}\left(k_m\right), \tag{20}$$

where $\{\alpha, \beta, \gamma\}$ must satisfy $\alpha + \beta + \gamma = 1$ since the ratio should be kept constant. In this study, we adopt $\{\alpha, \beta, \gamma\} = \{1/3,\ 1/3,\ 1/3\}$ for balancing all terms of cost equally, but we can arbitrarily adjust according to the environment, e.g., social norms or robot attributes.

C. TIME DELAY TO CHANGE DIRECTION k_m

We here define k_m and k_{mT}. Fig. 7 shows the process to determine k_m, where humans turn the direction from 15° to −15°, as an example. Humans take a couple of seconds to complete the turn. We first determined the time of starting a turn ($t = 0$), regarded as the time of passing through a turning point for simplicity, as shown in Figs. 9 (b) and (c). We then determined the time of finishing the turn k_e, regarded as the time when a human walks the target trajectory $\pm\theta_m$ [°] during a certain period s_m [s] to confirm the end of a turn. Finally, we calculate k_m (= $k_e - s_m$) as the time when the direction first enters within θ_m, as the minimum delay of the system. We set $\theta_m = 5°$ and $s_m = 0.1$ s for the experimental conditions in this study.

Then, we define k_{mT}, which is difficult to define since it should vary with the amount of turning direction, walking speed, and so on. Reference [34] reported that the time for humans to take a step is 0.5 s, and [35] reported that the upper limit of the pelvic rotation angle when humans turn is 30°, which means that a turn of 60° takes 1.0 s (two steps). In this study, we thus estimate k_{mT} by using proportional interpolation of the angle difference Δv_θ from the current to target trajectories when turning. k_{mT} is thus given by

$$k_{mT} = 0.5 \cdot \Delta v_\theta / 30. \tag{21}$$

Specifications of robot	
Movement Performance Platform	Velocity 1 m/s Acceleration 3 m/s² Omnidirectional movement by four omni wheels
Weight	32 kg

Specifications of LRF	
Model	UTM-30LX HOKUYO
Distance range	0.06–10 m, Max 30 m
Angular range	270°
Angular resolution	0.25°
Measurement resolution	±30 mm
Scanning time	25 ms/scan

FIGURE 8. Specification of robot equipped with omni wheels and LRF.

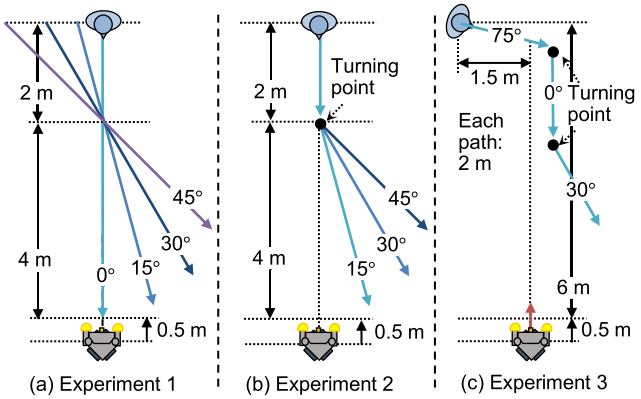

FIGURE 9. Overview of (a) Experiment 1 (straight walking), (b) Experiment 2 (single turn), and (c) Experiment 3 (zigzag walking with robot moving).

Note that the speed variables are not included in (21) since humans are required to walk naturally at a constant speed in the experiments.

VI. EVALUATION EXPERIMENTS

The robot and LRF specifications we used in this study are shown in Fig. 8. The experimental conditions and scenes are shown in Figs. 9 and 10, respectively. This study was approved by the Ethics Review Committee on Research with Human Subjects of Waseda University. We obtained written informed consent from each participant in this study.

A. DISPLACEMENT AND SYSTEM NOISE MODEL

To clarify the relationship between accuracy and latency, we first performed Exp. 1 to model the relationship between $d_e(k)$ and $\sigma_\epsilon^2(k)|_e$ with 11 participants aged 21–30. The participants walked straight with four different angles to the LRF (the robot), as shown in Fig. 9 (a). We asked the participants to walk at 1.0 m/s. Note that the ground truth of the velocity was guaranteed by practicing constant-speed walking before the experiments started. We collected point clouds and then calculated $d_e(k)$ and $\sigma_\epsilon^2(k)|_e$ using the five different $e = \{0.2, 0.4, 0.6, 0.8, 1.0\}$, which are enough to extract the trends of the results. The tolerance latency of 0.5 s would be acceptable for social robots since humans take 0.5 s for one step during a walk [34], so we set the maximum

FIGURE 10. Experimental scenes (Experiment 3). Participants walk by referring to marks on ground indicating target path.

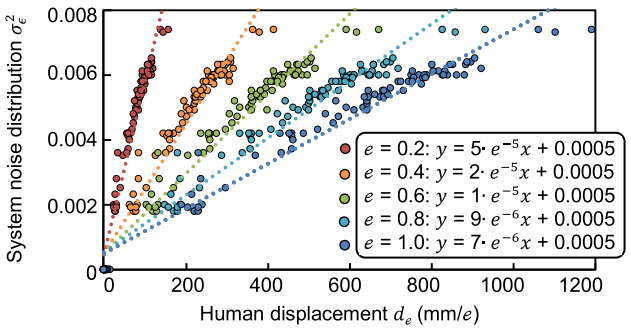

FIGURE 11. Results of Experiment 1: Relationship between human displacement d_e and system noise distribution σ_ϵ^2.

tolerable latency to 0.5 s and searched for the minimum $\sigma_\epsilon^2(k)|_e$ within 0.5 s. Fig. 11 shows the relationship between $d_e\,(k)$ and $\sigma_\epsilon^2(k)|_e$ for each e. We found that $d_e\,(k)$ and $\sigma_\epsilon^2(k)|_e$ have a proportional relationship, as

$$\sigma_\epsilon^2\,(k)\,|_e = a \cdot d_e\,(k) + b, \tag{22}$$

where a and b are the constant values shown in the lower-right part of Fig. 11. All plots include the results of 11 participants and four trajectories, so we can regard the equations as capturing general tendencies. The experimental results show that updating $\sigma_\epsilon^2(k)|_e$ according to observed human displacement $d_e\,(k)$ has the potential to enable the VES to adjust estimation accuracy and latency.

B. WINDOW SIZE ANALYSIS USING COSTS

Based on the displacement and system noise model, we analyzed the relationship among the window size $\{e,\ \ell\}$ and three costs. In Exp. 2, we prepared three different trajectories with a turning point (15°, 30°, and 45°) in the middle of the path (Fig. 9 (b)). The other experimental conditions were the same as Exp. 1. We calculated each cost using 25 $\{e \times \ell\}$ patterns $=$ {(0.2, 0.4, 0.6, 0.8, 1.0) $\times$ (0.2, 0.4, 0.6, 0.8, 1.0)} by (17)–(19). As Fig. 12 shows, we found that the three costs have different tendencies and that the cost maps between accuracy ($\mathcal{C}\left(\hat{\boldsymbol{v}}_\theta\right)$ and $\mathcal{C}\left(\hat{\boldsymbol{v}}_r\right)$) and latency ($\mathcal{C}\,(k_m)$) have a roughly inversed relationship, i.e., larger ℓ and smaller e lead to high accuracy while smaller ℓ leads to low latency. The results also show that $\mathcal{C}\left(\hat{\boldsymbol{v}}_\theta\right)$, $\mathcal{C}\left(\hat{\boldsymbol{v}}_r\right)$, and $\mathcal{C}\,(k_m)$ were the smallest in $\{e,\ \ell\} = \{0.2,\ 0.6\}$, $\{0.2,\ 1.0\}$, and $\{0.2,\ 0.4\}$, respectively, as indicated by the red dots in Fig. 12. When applying $\{\alpha,\ \beta,\ \gamma\} = \{1/3,\ 1/3,\ 1/3\}$ as explained before, the total cost $\mathcal{C}$ of $\hat{\boldsymbol{\mathcal{V}}}|_{e,\ell}$ was the lowest in $\{e, \ell\} = \{0.2,\ 0.6\}$. These findings confirm that window sizes can be used to control the estimation accuracy and latency.

C. COMPARISONS IN DIFFERENT TRAJECTORIES

To evaluate the applicability of the proposed system, we prepared more complex conditions with Exp. 3. A pedestrian changes its walking direction several times while the robot (LRF) moves (Fig. 9 (c)). The robot's trajectory is a straight line because we here inspect the impact of vibration noises due to the robot's movement. The human's walking direction and distance are in order of 75°(2 m), 0° (2 m), and 45° (2 m). The robot moves straight at 0.4 m/s. We asked the same 11 participants from Exps. 1 and 2 to walk at 1.0 m/s. Fig. 13 shows the three cost maps, indicating a trade-off relationship between accuracy and latency, the same as in Exp. 2. Compared with Exp. 2, $\mathcal{C}\left(\hat{\boldsymbol{v}}_r\right)$ was almost exactly the same, $\mathcal{C}\left(\hat{\boldsymbol{v}}_\theta\right)$ was similar, and $\mathcal{C}\,(k_m)$ was largely deviated. The deviated $\mathcal{C}\,(k_m)$ occurred because k_m would not converge when participants frequently changed their behaviors in a short time. This indicates that we need to redefine k_m in the future. $\mathcal{C}\left(\hat{\boldsymbol{v}}_\theta\right)$, $\mathcal{C}\left(\hat{\boldsymbol{v}}_r\right)$, and $\mathcal{C}\,(k_m)$ were the smallest in $\{e,\ \ell\} = \{0.2,\ 0.8\}$, $\{0.4,\ 0.8\}$, and $\{0.4,\ 0.4\}$, respectively, as indicated by the red dots in Fig. 13. When applying $\{\alpha,\ \beta,\ \gamma\} = \{1/3,\ 1/3,\ 1/3\}$, the total cost $\mathcal{C}$ of $\hat{\boldsymbol{\mathcal{V}}}|_{e,\ell}$ was the lowest in $\{e, \ell\} = \{0.2,\ 0.8\}$.

In Exps. 2 and 3, the shapes among the three cost maps are different. Moreover, the shapes of each cost map in Exp. 2 are also different from those in Exp. 3. Our above analysis confirms that the proposed VES with AWS can extract different characteristics to optimize accuracy and latency depending on walking trajectories.

D. PERFORMANCE OF VES WITH AWS

To evaluate the performance of controlling accuracy and latency in the proposed VES with AWS, we calculated the RSME of $\hat{\boldsymbol{v}}_\theta, \hat{\boldsymbol{v}}_r$, and k_m for Exps. 2 and 3, as shown in Fig. 14. We compared the RMSEs calculated using individually optimized parameters (direction, magnitude, and time delay), mean of 25 $\{e \times \ell\}$ patterns, parameters with the maximum cost, and equally balanced parameters. Using individually optimized parameters enhanced the target elements but degraded others, i.e., accuracy vs. latency. In contrast, the selection of equally balanced parameters ($\{e, \ell\} = \{0.2, 0.6\}$ for Exp. 2 and $\{0.2, 0.8\}$ for Exp. 3) led to all three elements achieving a high performance.

Here, we analyzed the effect of the equally balanced parameters (EBP). The EBP reduced the velocity direction RMSE to 1.86° and 5.11° for Exps. 2 and 3, respectively. In Exp. 3, the walking trajectory changes several times, and the robot's movement generates noises in point clouds, which made the direction RSME in Exp. 3 larger than that in Exp. 2, but the VES could select optimal window sizes $\{e, \ell\}$ that enable the direction RSME to minimize. Moreover, the EBP reduced the velocity magnitude RMSE to 0.0464 m/s and 0.146 m/s for Exps. 2 and 3, respectively. Even in Exp. 3,

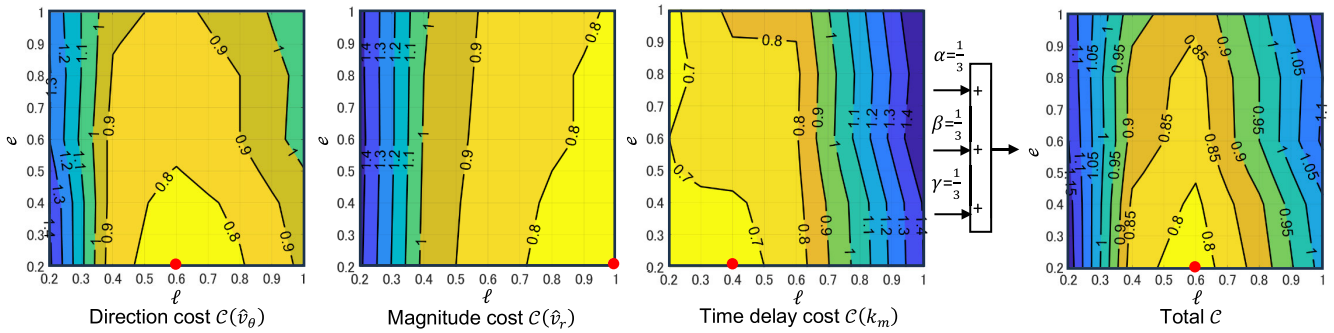

FIGURE 12. Results of Experiment 2. Three costs (direction, magnitude, time delay) are calculated in 25 patterns of (e, ℓ), and the total cost is calculated using $\{\alpha, \beta, \gamma\} = \{1/3, 1/3, 1/3\}$. Results show that ($e$, ℓ) = (0.2, 0.6) is the lowest cost (0.77). Red dots are (e, ℓ) with the lowest cost.

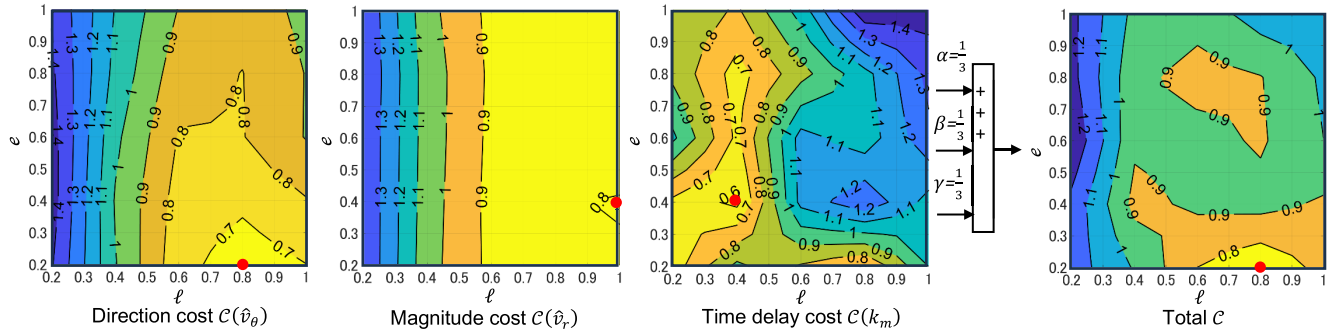

FIGURE 13. Results of Experiment 3. Three costs (direction, magnitude, time delay) are calculated in 25 patterns of (e, ℓ), and the total cost is calculated using $\{\alpha, \beta, \gamma\} = \{1/3, 1/3, 1/3\}$. Results show that ($e$, ℓ) = (0.2, 0.8) is the lowest cost (0.72). Red dots are (e, ℓ) with the lowest cost.

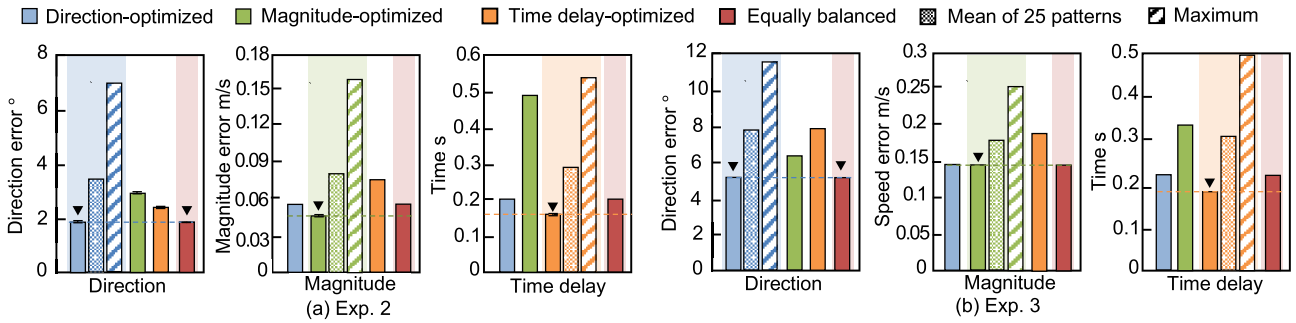

FIGURE 14. Performance of velocity estimation system with adjustable window size, including individually optimized parameters (direction, magnitude, and time delay), mean of 25 {$e \times \ell$} patterns, parameters with maximum cost, and equally balanced parameters for Exps. 2 and 3. Individually optimized parameters enhanced the target performance but degraded others. In contrast, equally balanced parameters achieved all three criteria simultaneously.

the VES could accurately estimate a velocity. Furthermore, the EBP decreased the time delay RMSE to 0.201 s and 0.218 s for Exps. 2 and 3, respectively. These values meet the requirement of being smaller than the time required for a human to take one step (i.e., 0.5 s) [34].

We next statistically analyzed the RSMEs using different optimized methods, including the three individually optimized parameters (direction, magnitude, and time delay) and equally balanced parameters. We adopted the Friedman test for multiple comparisons of three element's RMSE. Then, we applied Scheffe's paired comparison of different optimized methods as a post-hoc test. The results of the statistical test are summarized in Table 2. The Friedman test revealed statistical differences in all conditions. Moreover, Scheffe's paired comparison revealed that using individually optimized parameters, the VES could statistically improve the target elements than non-target elements for Exps. 2 and 3.

The above analysis confirmed that the proposed VES with AWS had a higher performance corresponding to the target purpose than when the VES selected other parameters.

E. OPTIMAL PARAMETERS FOR INDIVIDUALS

Finally, we investigate individual differences among participants. Fig. 15 shows selected parameters in the direction,

TABLE 2. Statistical analysis of three element's RSME for different optimized methods in Exps. 2 and 3.

Element		Optimized method		
	Friedman test	Scheffe's paired comparison*		
Exp. 2		D(B)–M	D(B)–T	M–T
Direction	$\chi^2(2) = 25.88$, $p<0.001$	$p<0.001$	$p<0.005$	n.s.
Magnitude	$\chi^2(2) = 56.18$, $p<0.001$	$p<0.05$	$p<0.001$	$p<0.001$
Time delay	$\chi^2(2) = 44.44$, $p<0.001$	$p<0.001$	n.s.	$p<0.001$
Exp. 3		D(B)–M	D(B)–T	M–T
Direction	$\chi^2(2) = 13.27$, $p<0.005$	n.s.	$p<0.005$	n.s.
Magnitude	$\chi^2(2) = 16.55$, $p<0.001$	$p<0.005$	$p<0.005$	$p<0.005$
Time delay	$\chi^2(2) = 12.88$, $p<0.005$	$p<0.05$	n.s.	$p<0.005$

* D: Direction, M: Magnitude, T: Time delay, B: Balanced

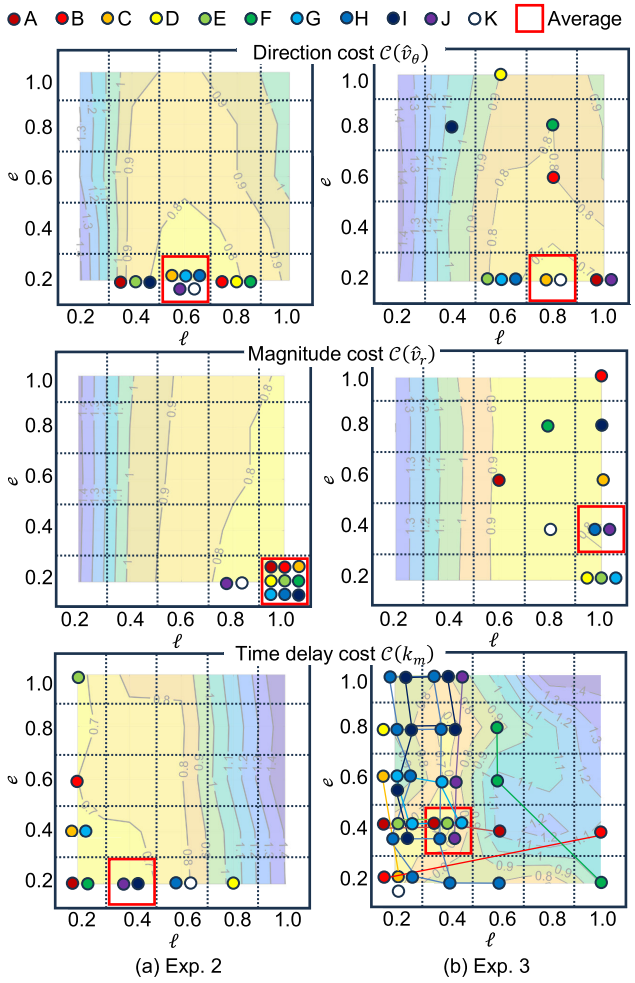

FIGURE 15. Parameters with lowest cost for direction, magnitude, and time delay for each pedestrian for (a) Exp. 2 and (b) Exp. 3. The best parameters to minimize estimation error or time delay are different, and these parameters also differ among individuals. Several parameters had the minimum cost.

magnitude, and time delay for the 11 participants in Exps. 2 and 3. As shown in Fig. 15 (a), the optimal parameters in Exp. 2 were similar among individuals, and in particular, the parameters in the magnitude were almost identical ($\{e, \ell\} = \{0.2, 1.0\}$). The walking trajectory in Exp. 2 was straight, so no individual differences appeared. However, the parameters for the direction are spread along the ℓ axis, meaning that the VES could deduce that lateral sway during walking differed among participants. The parameters for the time delay were slightly deviated. As shown in Fig. 15 (b), optimal parameters in Exp. 3 were more deviated than Exp. 2, and those for time delay were remarkable. These results imply that the way of turning, e.g., turning radius and angular velocity, deviates largely depending on the participant. We confirmed from this analysis that the VES with AWS could reveal the existence of optimal window sizes suitable for individuals.

F. LIMITATION AND FURTHER IMPROVEMENTS

Through this study, we found that the proposed framework was capable of identifying window sizes suitable for walking trajectories and individual gaits for time-series-based velocity estimation and that the proposed VES with AWS could improve the estimation performance following a target purpose. The newly proposed cost function was useful for controlling estimation accuracy and latency. Moreover, the proposed system could be applied to a use case where an LRF is installed on a mobile robot. The preliminary results of this study would contribute to developing individually optimized measurements using data based on the robot's movement experience. However, there are some limitations to be addressed in the future, as follows.

1) AUTOMATIC PARAMETER TUNING SYSTEM

In this study, the weight coefficients $\{\alpha, \beta, \gamma\}$ are temporally set to 1/3 (as described in subsection V-B). However, these coefficients should be tuned in accordance with the purpose or situation, e.g., when a robot needs to plan an accurate trajectory by suppressing the human's lateral sway, overtake a human by suppressing the human's longitudinal sway, or immediately avoid a human who may be approaching from out of sight. Moreover, the proposed VES currently learns the window sizes suitable for a condition beforehand, but the robot should be able to adjust window sizes autonomously in the introduced environments [36]. We plan to develop an automatic parameter tuning system to tune the above coefficient in real time while combining error-tolerant navigation [17] if the system makes a judgment error.

2) DEFINITION OF TIME DELAY

In Exp. 3, $\mathcal{C}(k_m)$ deviated largely because participants frequently changed their behaviors in a short time, so we need to modify how to define k_m and k_{mT}. Specifically, we will observe walking behaviors using a skeleton-based gait model when turning and introduce a system to predict a pedestrian's motion using other obtainable information, such as head movement and gaze.

VII. CONCLUSION AND FUTURE WORK

In this paper, we proposed a new approach to estimating human walking velocity with a function to control accuracy and latency. The proposed velocity estimation system (VES)

consists of a Kalman filter (KF), least squares (LS), and a window-size adjustable system. The adjustable system tunes a window size e for calculating a system noise in KF and a window size ℓ for calculating a velocity in LS by using a cost function including accuracy (velocity angle and magnitude) and latency (time delay) that can learn $\{e, \ell\}$ suitable for an intended purpose (high accuracy, low latency, or a balance between the two). Experiments in which a robot was moving and humans were continuously changing their behaviors demonstrated that the proposed VES significantly improved the measurement performance thanks to selecting the appropriate window sizes compared to using other parameters. Our proposed method adopts a relative improvement strategy by optimizing system parameters, such as [37], not an absolute improvement strategy by collecting mathematical models, so it can be applied to other estimation methods that handle time-series data as a relative improvement approach. This study will contribute to developing a framework for a VES that can enable measurement performances suitable for the target conditions and enhance the performance of mobile robot navigation.

In future work, to achieve high-accuracy and low-latency estimation in a real unstructured environment, we will modify the cost function for finding the optimal window-size parameters, create a model to automatically update the weight coefficients, and perform experiments under a variety of conditions to evaluate and improve our proposed system.

REFERENCES

[1] S.-Y. Chung and H.-P. Huang, "A mobile robot that understands pedestrian spatial behaviors," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2010, pp. 5861–5866.

[2] M. Kamezaki, Y. Tsuburaya, T. Kanada, M. Hirayama, and S. Sugano, "Reactive, proactive, and inducible proximal crowd robot navigation method based on inducible social force model," *IEEE Robot. Autom. Lett.*, vol. 7, no. 2, pp. 3922–3929, Apr. 2022.

[3] M. Kamezaki, A. Kobayashi, Y. Yokoyama, H. Yanagawa, M. Shrestha, and S. Sugano, "A preliminary study of interactive navigation framework with situation-adaptive multimodal inducement: Pass-by scenario," *Int. J. Social Robot.*, vol. 12, no. 2, pp. 567–588, May 2020.

[4] M. C. Shrestha, T. Onishi, A. Kobayashi, M. Kamezaki, and S. Sugano, "Communicating directional intent in robot navigation using projection indicators," in *Proc. 27th IEEE Int. Symp. Robot Human Interact. Commun. (RO-MAN)*, Aug. 2018, pp. 746–751.

[5] K. Guo, D. Wang, T. Fan, and J. Pan, "VR-ORCA: Variable responsibility optimal reciprocal collision avoidance," *IEEE Robot. Autom. Lett.*, vol. 6, no. 3, pp. 4520–4527, Jul. 2021.

[6] D. Wilkie, J. van den Berg, and D. Manocha, "Generalized velocity obstacles," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2009, pp. 5573–5578.

[7] A. Bera, T. Randhavane, R. Prinja, and D. Manocha, "SocioSense: Robot navigation amongst pedestrians with social and psychological constraints," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 7018–7025.

[8] Q. Yao, Z. Zheng, L. Qi, H. Yuan, X. Guo, M. Zhao, Z. Liu, and T. Yang, "Path planning method with improved artificial potential field—A reinforcement learning perspective," *IEEE Access*, vol. 8, pp. 135513–135523, 2020.

[9] M. Kamezaki, A. Kobayashi, R. Kono, M. Hirayama, and S. Sugano, "Dynamic waypoint navigation: Model-based adaptive trajectory planner for human-symbiotic mobile robots," *IEEE Access*, vol. 10, pp. 81546–81555, 2022.

[10] G. Whelch and G. Bishop, "An introduction to the Kalman filter," Dept. Comput. Sci., Univ. North Carolina, Chapel Hill, NC, USA, Tech. Rep. TR95-041, 2006.

[11] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, Mar. 1960.

[12] K. Lin, Z. Guo, F. Yang, J. Huang, and Y. Zhang, "Kalman filter-based multi-object tracking algorithm by collaborative multi-feature," in *Proc. IEEE 2nd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Mar. 2017, pp. 1239–1244.

[13] Q. Xiao, F. Sun, R. Ge, K. Chen, and B. Wang, "Human tracking and following of mobile robot with a laser scanner," in *Proc. 2nd Int. Conf. Adv. Robot. Mechatronics (ICARM)*, Aug. 2017, pp. 675–680.

[14] E.-J. Jung, J. H. Lee, B.-J. Yi, J. Park, S. Yuta, and S.-T. Noh, "Development of a laser-range-finder-based human tracking and control algorithm for a marathoner service robot," *IEEE/ASME Trans. Mechatronics*, vol. 19, no. 6, pp. 1963–1976, Dec. 2014.

[15] A. H. Adiwahono, V. B. Saputra, K. P. Ng, W. Gao, Q. Ren, B. H. Tan, and T. W. Chang, "Human tracking and following in dynamic environment for service robots," in *Proc. IEEE Region 10 Conf.*, Nov. 2017, pp. 3068–3073.

[16] H. Iida and T. Yamamuro, "Kinetic analysis of the center of gravity of the human body in normal and pathological gaits," *J. Biomechanics*, vol. 20, no. 10, pp. 987–995, Jan. 1987.

[17] M. Kamezaki, R. Kono, A. Kobayashi, H. Yanagawa, T. Onishi, Y. Tsuburaya, M. Shrestha, and S. Sugano, "Toward error-tolerant robot navigation: Sequential inducement based on intent conveyance from robot to human and its achievement," *Int. J. Social Robot.*, vol. 15, no. 2, pp. 297–316, Feb. 2023.

[18] J. Hoon Lee, K. Abe, T. Tsubouchi, R. Ichinose, Y. Hosoda, and K. Ohba, "Collision-free navigation based on people tracking algorithm with biped walking model," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2008, pp. 2983–2989.

[19] S. Cho, C. Kim, M. Sunwoo, and K. Jo, "Robust localization in map changing environments based on hierarchical approach of sliding window optimization and filtering," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 4, pp. 3783–3789, Apr. 2022.

[20] J. Carpentier, M. Benallegue, and J.-P. Laumond, "On the centre of mass motion in human walking," *Int. J. Autom. Comput.*, vol. 14, no. 5, pp. 542–551, Oct. 2017.

[21] V. T. Inman, H. J. Ralston, and F. Todd, *Human Walking*. Baltimore, MD, USA: Williams & Wilkins, 1981.

[22] E. Foxlin, "Pedestrian tracking with shoe-mounted inertial sensors," *IEEE Comput. Graph. Appl.*, vol. 25, no. 6, pp. 38–46, Nov. 2005.

[23] S. Koo and D.-S. Kwon, "Recognizing human intentional actions from the relative movements between human and robot," in *Proc. RO-MAN - 18th IEEE Int. Symp. Robot Hum. Interact. Commun.*, Sep. 2009, pp. 939–944.

[24] H. T. Duong and Y. S. Suh, "Walking parameters estimation based on a wrist-mounted inertial sensor for a Walker user," *IEEE Sensors J.*, vol. 17, no. 7, pp. 2100–2108, Apr. 2017.

[25] A. Laudanski, S. Yang, and Q. Li, "A concurrent comparison of inertia sensor-based walking speed estimation methods," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2011, pp. 3484–3487.

[26] K. Brzostowski, "Toward the unaided estimation of human walking speed based on sparse modeling," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 6, pp. 1389–1398, Jun. 2018.

[27] M. Bertozzi, A. Broggi, A. Fascioli, A. Tibaldi, R. Chapuis, and F. Chausse, "Pedestrian localization and tracking systen with Kalman filtering," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2004, pp. 584–589.

[28] Q. Xu, H. Wu, J. Wang, H. Xiong, J. Liu, and K. Li, "Roadside pedestrian motion prediction using Bayesian methods and particle filter," *IET Intell. Transp. Syst.*, vol. 15, no. 9, pp. 1167–1182, Sep. 2021.

[29] T. Hunter, P. Abbeel, and A. Bayen, "The path inference filter: Model-based low-latency map matching of probe vehicle data," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 2, pp. 507–529, Apr. 2014.

[30] Q. Ren, Q. Zhao, H. Qi, and L. Li, "Real-time target tracking system for person-following robot," in *Proc. 35th Chin. Control Conf. (CCC)*, Jul. 2016, pp. 6160–6165.

[31] P. Ratsamee, Y. Mae, K. Ohara, T. Takubo, and T. Arai, "People tracking with body pose estimation for human path prediction," in *Proc. IEEE Int. Conf. Mechatronics Autom.*, Aug. 2012, pp. 1915–1920.

[32] E. Machida, M. Cao, T. Murao, and H. Hashimoto, "Human motion tracking of mobile robot with Kinect 3D sensor," in *Proc. SICE Annu. Conf. (SICE)*, Aug. 2012, pp. 2207–2211.

[33] A. Björck, *Numerical Methods for Least Squares Problems*. Philadelphia, PA, USA: SIAM, 1996.

[34] K. E. Peyer, C. A. Brassey, K. A. Rose, and W. I. Sellers, "Locomotion pattern and foot pressure adjustments during gentle turns in healthy subjects," *J. Biomechanics*, vol. 60, pp. 65–71, Jul. 2017.
[35] Y. Akiyama, H. Toda, T. Ogura, S. Okamoto, and Y. Yamada, "Classification and analysis of the natural corner curving motion of humans based on gait motion," *Gait Posture*, vol. 60, pp. 15–21, Feb. 2018.
[36] M. Kamezaki, H. Iwata, and S. Sugano, "Basic input–output gain tuning system based on control input histogram leveling for human-operated machines," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 6, pp. 2857–2869, Dec. 2021.
[37] M. Kamezaki, H. Iwata, and S. Sugano, "Relative accuracy enhancement system based on internal error range estimation for external force measurement in construction manipulator," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2011, pp. 3734–3739.

MITSUHIRO KAMEZAKI (Member, IEEE) received the B.S., M.S., and Dr.Eng. degrees in mechanical engineering from Waseda University, Tokyo, Japan, in 2005, 2007, and 2010, respectively. From 2010 to 2013, he was a Research Associate with the Department of Modern Mechanical Engineering, Waseda University. He was an Assistant Professor and an Associate Professor with the Research Institute for Science and Engineering, Waseda University, from 2013 to 2018 and 2018 to 2023, respectively. Currently, he is a Project Professor with the Graduate School of Engineering, The University of Tokyo, Tokyo. His current research interests include intelligent mobile robot, human-aware interactive navigation, personal mobility devices, and operator support system in operated-work machines. He is a member of the IEEE RAS, IEEE IES, SICE, RSJ, and JSME. He received the ICROS Best Paper Award Finalist from the IEEE/RSJ IROS 2011, the SICE Young Author's Award Finalist from the 2013 SICE Annual Conference, the Young Investigation Excellence Award from RSJ, in 2016, the Best Paper Award from the IEEE/ASME AIM 2016, and the IEEE Robotics and Automation Letters Best Paper Award, in 2022.

MICHIAKI HIRAYAMA received the B.S. and M.S. degrees in modern mechanical engineering from Waseda University, Tokyo, Japan, in 2018 and 2021, respectively. His research interests include dynamic path planning and self-parameter tunable system for autonomous mobile robots.

RYOSUKE KONO received the B.S. and M.S. degrees in modern mechanical engineering from Waseda University, Tokyo, Japan, in 2017 and 2019, respectively. His research interests include error-tolerant navigation and robust human-movement estimation for autonomous mobile robots.

YUSUKE TSUBURAYA received the B.S. and M.S. degrees in modern mechanical engineering from Waseda University, Tokyo, Japan, in 2017 and 2019, respectively. His research interests include in-crowd navigation using social force model and physical interaction during navigation for autonomous mobile robots.

SHIGEKI SUGANO (Fellow, IEEE) received the B.S., M.S., and Dr.Eng. degrees in mechanical engineering from Waseda University, Tokyo, Japan, in 1981, 1983, and 1989, respectively. Since 1986, he has been a Faculty Member with the Department of Mechanical Engineering, Waseda University, where he is currently a Professor. Since 2014, he has been the Dean of the School/Graduate School of Creative Science and Engineering, Waseda University. His research interests include human-symbiotic anthropomorphic robot design, dexterous and safe manipulator design, and human–robot communication. He is a Fellow of four academic societies, Japan Society of Mechanical Engineers (JSME), the Society of Instrument and Control Engineers (SICE), and the Robotics Society of Japan (RSJ). He served as the General Chair for the IEEE/ASME International Conference on Advanced Intelligent Mechatronics, in 2003, and the IEEE/RSJ International Conference on Intelligent Robots and Systems, in 2013. From 2001 to 2010, he served as the President for Japan Association for Automation Advancement. He was the General Co-Chair of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2006) and the Program Co-Chair of the 2009 IEEE International Conference on Robotics and Automation (ICRA2009). He also served as the General Co-Chair for the 2012 IEEE International Conference on Robotics and Automation (ICRA2012) and the Program Chair for the 2012 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM2012). In 2017, he served as the President for SICE.