Hierarchical Cooperative Tracking of Vehicles and People Using Laser Scanners Mounted on Multiple Mobile Robots

Yuto Tamura, Ryohei Murabayashi
Graduate School of Science and Engineering
Doshisha University
Kyotanabe, Kyoto 610-0394 Japan
e-mail: {ytamura915, ryo040978}@gmail.com

Masafumi Hashimoto, Kazuhiko Takahashi
Faculty of Science and Engineering
Doshisha University
Kyotanabe, Kyoto 610-0394 Japan
e-mail: {mhashimo, katakaha}@mail.doshisha.ac.jp

***Abstract*—This paper presents a tracking (estimation of the pose and size) of moving objects such as pedestrians, cars, motorcycles, and bicycles, using multiple mobile robots as sensor nodes. In this cooperative-tracking method, nearby sensor nodes share their tracking information, enabling the tracking of objects that are invisible or partially visible to an individual sensor node. The cooperative-tacking method can then make the tracking system more reliable than the conventional individual tracking method by a single robot. We previously presented a centralized architecture of cooperative tracking. Each sensor node detected moving objects in its own laser-scanned images captured by a single-layer laser scanner. It then sent measurement information related to the moving objects to a central server. The central server estimated the objects' poses (positions and velocities) and sizes from the measurement information using Bayesian filter. However, such a centralized method might have poor dependability and impose a computational burden upon the central server. To address this problem, this paper presents hierarchical architecture of cooperative tracking. Each sensor node locally estimates the poses and sizes of moving objects and then sends these estimates to the central server, which then merges the pose and size estimates. Experimental results using two sensor nodes in outdoor environments show that the proposed hierarchical cooperative-tracking method provides slightly inferior tracking accuracy and has a smaller computational cost in the central server than a previous centralized method.**

Keywords—moving-object tracking; cooperative tracking; centralized and hierarchical methods; laser scanner; mobile robot.

I. Introduction

This paper is an extended version of an earlier paper presented at the IARIA Conference on Intelligent Systems and Applications (INTELLI 2016) [1] in Barcelona. Throughout this paper, the term 'tracking' means the estimation of the pose (position and velocity) and size of a moving object.

The tracking of multiple moving objects (such as people, cars, and bicycles) in environments is an important issue in the safe navigation of mobile robots and vehicles. The use of vision, radar or laser scanner (Lidar) in mobile robotics and vehicle automation has attracted considerable interest [2]–[8]. When compared with vision-based tracking, laser-based tracking is insensitive to lighting conditions and requires less data processing time. Furthermore, due to its directionality, laser-based tracking provides better tracking accuracy than radar-based tracking. Therefore, in this paper, we focus on a tracking method for moving objects using laser scanners mounted on mobile robots and vehicles.

Many studies have been conducted on multi-robot coordination and cooperation [9][10]. When multiple robots are located in the same vicinity, they can share their sensing data through a communication network system. Thus, the multi-robot team can be considered to be a multi-sensor system. Even if moving objects are located outside the sensing area of a robot or are occluded in crowded environments, they can be recognized using tracking data from other nearby robots. Hence, multi-robot system can improve the accuracy and reliability with which moving objects are tracked.

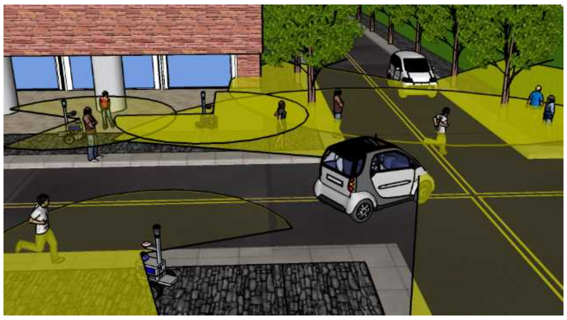


Figure 1. Application of cooperative tracking to a vehicle automation field.

Such cooperative tracking using multiple robots and vehicles can also be applied to vehicle automation, including intelligent transportation systems (ITS) and systems for personal mobility devices, as shown in Fig. 1. Cooperative tracking enables the detection of moving objects in the blind-spot of each vehicle and can be used to detect sudden changes in a crowded urban environment such as people appearing on roads or vehicles making unsafe lane changes. It can therefore prevent traffic accidents.

In this paper, we present a hierarchical cooperative-tracking method for moving objects using multiple mobile robots as sensor nodes. The sensor nodes locally track moving objects and transmit the tracking information to a central server, which then merges the tracking information.

For simplicity, in this paper, moving-object tracking using multiple mobile sensor nodes is referred to as 'cooperative tracking,' whereas that by an individual robot in a team is referred to as 'individual tracking.'

The rest of the paper is organized as follows. Section II presents an overview of related work. Section III gives our experimental system. In Sections IV to VII, cooperative tracking method is discussed. In Section VIII, we describe experiments of moving-object tracking using two sensor nodes in outdoor environments. We will present our conclusions in Section IX.

II. Related Work

We previously presented a cooperative people-tracking method in which multiple mobile robots and vehicles were used as mobile sensor nodes and equipped with laser scanners [11][12]. The covariance intersection method [13] was applied to operate the tracking system effectively in a decentralized manner without a central server. In cooperative people tracking, each person could be assumed to be a point due to their small size, and mass-point tracking (only pose estimation) was then performed.

However, in the real world, several types of moving objects exist, such as people, cars, bicycles, and motorcycles. Therefore, we should design a cooperative-tracking system for these moving objects. In vehicle (car, motorcycle, and bicycle) tracking, we have to consider moving objects as rigid bodies and estimate both the poses and sizes to avoid collisions in a crowded environment. Tracking of a rigid body is known as extended-object tracking, and many related studies have been conducted [14]–[18]. However, to the best of our knowledge, cooperative tracking using multiple mobile sensor nodes covers only mass-point tracking under the assumption that the tracked object is small. It estimates only the object's pose but does not estimate its size [19]–[25].

Therefore, we presented a laser-based cooperative-tracking method for rigid bodies that estimates both poses and sizes of people and vehicles using multiple mobile sensor nodes [26]. In a crowded environment, a vehicle can be occluded or only rendered partially visible to each sensor node. To correctly estimate the size of the vehicle, the laser measurements captured by sensor nodes in the team have to be merged. Our previous cooperative-tracking method for rigid bodies applied a centralized architecture. Each sensor node detected laser measurements related to the moving objects in its sensing area and transmitted the measurement information to a central server, which then estimated the poses and sizes of the objects. Such a centralized architecture imposes a computational burden upon the central server. Furthermore, the architecture has a weakness against fault in the communication system between sensor nodes and the central server.

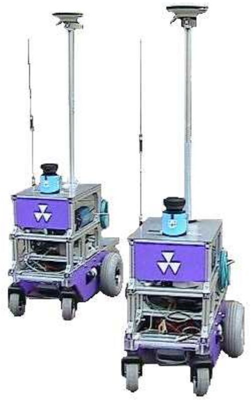


Figure 2. Overview of the mobile sensor nodes.

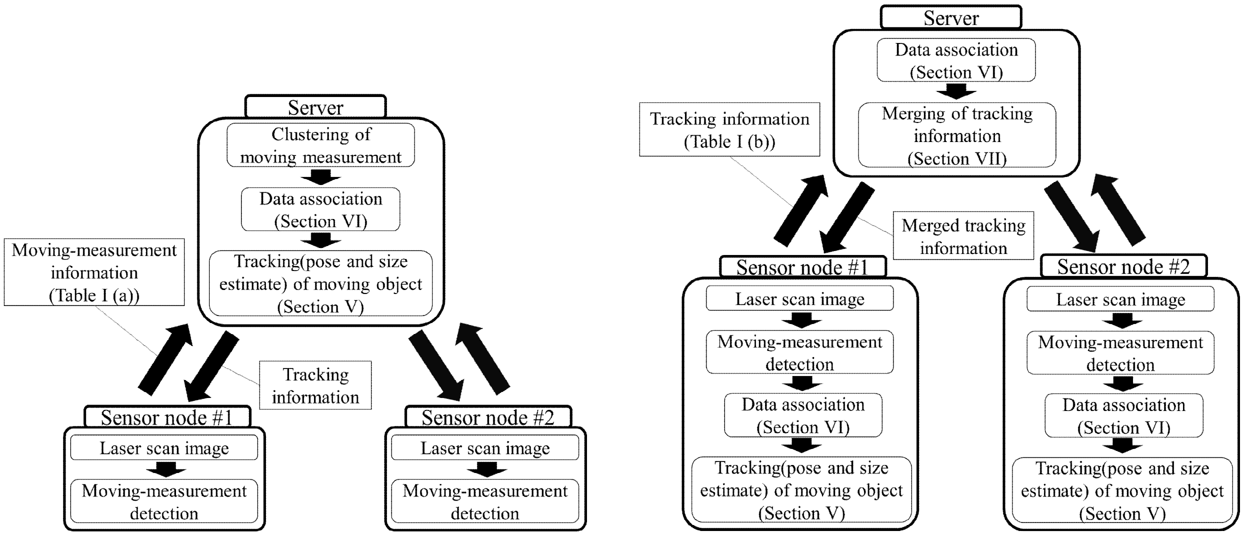
To address this problem, in this paper, we present a hierarchical method for cooperative tracking through which the poses and sizes of moving objects are locally estimated by the sensor nodes. Furthermore, these estimates are then merged by a central server. We will treat both vehicles and people as rigid bodies.

III. Experimental System

Fig. 2 shows the mobile-sensor node system used in our experiments. Each of the two sensor nodes has two independently driven wheels. A wheel encoder is attached to each drive wheel to measure its velocity. A yaw-rate gyro is attached to the chassis of each robot to sense the turning velocity. These internal sensors calculate the robot's pose using dead reckoning.

Each sensor node is equipped with a forward-looking laser scanner (SICK LMS100) to capture laser-scanned images that are represented by a sequence of distance samples in a horizontal plane with a field of view of 270°. The angular resolution of the laser scanner is 0.5°, and each scan image comprises 541 distance samples. Each sensor node is also equipped with RTK-GPS (Novatel ProPak-V3 GPS). The sampling frequency of all sensors is 10 Hz.

We use broadcast communication over a wireless local area network to exchange information between the central server and the sensor nodes. The model of the computers used in the sensor nodes and the central server is Iiyama 15X7100-i7-VGB with a 2.8 GHz Intel core i7-4810MQ processor, and the operating system is Microsoft Windows 7 Professional.


Server
Clustering of moving measurement
Data association (Section VI)
Tracking(pose and size estimate) of moving object (Section V)
Moving-measurement information (Table I (a))
Tracking information
Sensor node #1
Laser scan image
Moving-measurement detection
Sensor node #2
Laser scan image
Moving-measurement detection
Server
Data association (Section VI)
Merging of tracking information (Section VII)
Tracking information (Table I (b))
Merged tracking information
Sensor node #1
Laser scan image
Moving-measurement detection
Data association (Section VI)
Tracking(pose and size estimate) of moving object (Section V)
Sensor node #2
Laser scan image
Moving-measurement detection
Data association (Section VI)
Tracking(pose and size estimate) of moving object (Section V)

(a) Centralized method

(b) Hierarchical method

Figure 3. System overview of cooperative tracking.

IV. Overview of Cooperative-tracking System

In Section VIII, we will evaluate our hierarchical cooperative-tracking method through experiments by comparing it with our previous centralized method. To allow readers to better understand our cooperative-tracking method, we detail both the centralized and hierarchical methods of cooperative tracking. Figs. 3 (a) and (b) show a sequence of centralized and hierarchical cooperative-tracking, respectively, using two sensor nodes and a central server.

A. Centralized Method [26]

Each sensor node independently finds moving objects in its own laser-scanned image using an occupancy-grid method [27]. Laser measurements (positions) are mapped onto the grid map represented in the world coordinate frame Σ_w. The mapped measurements are classified into moving and static measurements using the occupancy-grid method. The moving measurements are considered to originate from moving objects, whereas the static measurements are considered to be from static objects. The sensor node uploads moving-measurement information to a central server. This information is detailed in Table I (a).

Moving measurements coming from the same moving object have similar positions, whereas those from different moving objects are significantly different. Thus, the central server clusters moving measurements sent from two sensor nodes by checking the gap between two adjacent measurements. Subsequently, the server estimates the poses and sizes of the moving objects using the methods to be presented in Sections V and VI. The estimated information is then fed back to the sensor nodes.

The cell size of the grid map is set at 0.3 × 0.3 m in our experiments. To map the laser-scanned images onto the grid map, each sensor node accurately identifies its own position and orientation in Σ_w based on dead reckoning and GPS information via an extended Kalman filter. Our moving measurement detection using the occupancy-grid method and self-localization using an extended Kalman filter are detailed in references [27] and [12], respectively.

B. Hierarchical Method

In the hierarchical method, each sensor node classifies the mapped laser measurements into moving and static

TABLE I. Information and Data Volume Sent from Each Sensor Node to Central Server

(a) Centralized Method

Information	Data volume [bit]
Time stamp	2 × 32
Pose of sensor node (x, y, θ)	3 × 32
The number of moving objects (n)	32
The number of moving measurements comprising each moving object (m)	n × 32
Coordinates of each moving measurement (x, y)	n × m × 2 × 32
Total	(6 + n + 2nm) × 32

(b) Hierarchical Method

Information	Data volume [bit]
Time stamp	2 × 32
The number of tracked objects (n)	32
Position and velocity estimate of each tracked object $(x, \dot{x}, y, \dot{y})$	n × 4 × 32
Heading of each tracked object (θ)	32
Size (width and lenglth) of each tracked object (W, L)	n × 2 × 32
Total	(4 + 6n) × 32

measurements using the occupancy-grid method and clusters moving measurements by checking the gap between two adjacent measurements. Subsequently, the sensor node locally estimates the poses and sizes of moving objects using the methods shown in Sections V and VI. This means individual tracking. The tracking information of the moving objects, which is detailed in Table I (b), is then uploaded to the central server.

After receiving the information regarding the tracked objects from two sensor nodes, the central server merges the information using the method in Section VII to improve the tracking accuracy. The merged poses and sizes of the moving objects are then fed back to the sensor nodes.

V. Pose and Size Estimation

In the hierarchical cooperative-tracking method, the pose and size are locally estimated by sensor nodes, whereas, in the centralized method, they are estimated by a central server.

We represent the shape of a moving object using a rectangle with width, *W* and length, *L*. We detail the size-estimation method in Fig. 4, where red circles indicate laser measurements of the moving object (hereafter referred to as moving measurements), and green lines are the feature lines extracted from these measurements. The green dashed rectangle is the estimated rectangle, and the green star is the centroid of that rectangle. As shown in Fig. 4, an $x_v y_v$-coordinate frame is defined, on which the y_v-axis aligns with the heading (orange arrow) of a tracked object. From clustered moving measurements, we extract the width, W_{meas} and length, L_{meas}.

When a moving object is perfectly visible, its size can be estimated from these moving measurements. In contrast, when it is partially occluded by other objects, its size cannot be accurately estimated. Therefore, the size of a partially visible object is estimated using the following equation [14]:

$$\begin{cases} W_{(t)} = W_{(t-1)} + G_W (W_{meas} - W_{(t-1)}) \\ L_{(t)} = L_{(t-1)} + G_L (L_{meas} - L_{(t-1)}) \end{cases} \tag{1}$$

where *W* and *L* are estimates of the width and length, respectively, and *t* and *t*-1 are time steps. *G* is the filter gain, given by $G = 1 - \sqrt[t]{(1-p)}$ [14], and *p* is a parameter. As the value of *p* increases, the reliabilities of the current measurements of W_{meas} and L_{meas} increase. We assume that a vehicle passes at 60 km/h in front of the sensor node. After the vehicle enters the surveillance area of the sensor node, we aim to estimate 99% of the size (*p* = 0.99) within 10 scans (1 s) of the laser scanner. We can then determine *G* as follows:

$$G = \begin{cases} 1 - \sqrt[t]{(1-0.99)} & \text{for } t \le 10 \\ 1 - \sqrt[10]{(1-0.99)} = 0.369 & \text{for } t > 10 \end{cases} \tag{2}$$

For a perfectly visible object, we set the gain G_w (G_L) as follows: if $W_{(t\text{-}1)}$ ($L_{(t\text{-}1)}$) < W_{meas} (L_{meas}), then G_w (G_L) =1, else G_w (G_L) = 0.

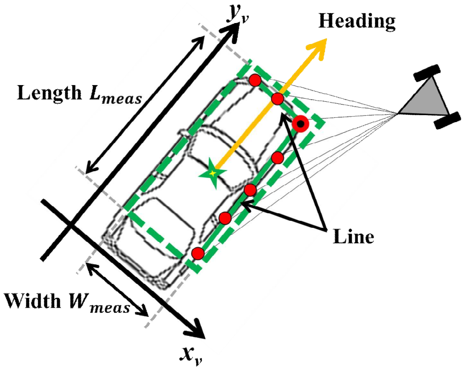


Figure 4. Size estimation of a vehicle.

The estimated size of the tracked object is used to classify the object as a person or a vehicle (i.e., car, motorcycle, and bicycle). If the estimated length or width is larger than 0.8 m, the object is determined to be a vehicle. If the size is less than 0.8 m, it is determined to be a person.

We then define the centroid position (green star in Fig. 4) of the rectangle estimated from (1) by (x, y) in Σ_w. From the centroid position, the pose of the tracked object in Σ_w is estimated using the Kalman filter [28] under the assumption that the object is moving at an almost constant velocity. The rate kinematics is given by:

$$\begin{aligned} \boldsymbol{x}_{(t)} &= \boldsymbol{F}\boldsymbol{x}_{(t-1)} + \boldsymbol{G}\Delta\boldsymbol{x}_{(t-1)} \\ &= \begin{pmatrix} 1 & \tau & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \tau \\ 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{x}_{(t-1)} + \begin{pmatrix} \tau^2/2 & 0 \\ \tau & 0 \\ 0 & \tau^2/2 \\ 0 & \tau \end{pmatrix} \Delta\boldsymbol{x}_{(t-1)} \end{aligned} \tag{3}$$

where $\boldsymbol{x} = (x, \dot{x}, y, \dot{y})^T$. $\Delta\boldsymbol{x} = (\Delta\ddot{x}, \Delta\ddot{y})^T$ is an unknown acceleration (plant noise). τ (=0.1 s) is the sampling period of the laser scanner.

The measurement model related to the moving object is then:

$$\begin{aligned} \boldsymbol{z}_{(t)} &= \boldsymbol{H}\boldsymbol{x}_{(t)} + \Delta\boldsymbol{z}_{(t)} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \boldsymbol{x}_{(t)} + \Delta\boldsymbol{z}_{(t)} \end{aligned} \tag{4}$$

where $z = (z_x, z_y)^T$ is the centroid position represented in Σ_w. Δz is the measurement noise.

From (3), the pose $\hat{x}$ of the object and its associated error covariance $\boldsymbol{P}$ can be predicted using the Kalman filter:

$$\begin{cases} \hat{\boldsymbol{x}}_{(t/t-1)} = \boldsymbol{F}\hat{\boldsymbol{x}}_{(t-1)} \\ \boldsymbol{P}_{(t/t-1)} = \boldsymbol{F}\boldsymbol{P}_{(t-1)}\boldsymbol{F}^T + \boldsymbol{G}\boldsymbol{Q}_{(t-1)}\boldsymbol{G}^T \end{cases} \tag{5}$$

where $\boldsymbol{Q}$ is the covariance of the plant noise $\Delta\boldsymbol{x}$.

When the measurement z is obtained from the tracked object, the pose of the tracked object and its associated error covariance are updated using:

$$\begin{cases} \hat{\boldsymbol{x}}_{(t)} = \hat{\boldsymbol{x}}_{(t/t-1)} + \boldsymbol{K}_{(t)}(\boldsymbol{z}_{(t)} - \boldsymbol{H}_{(t)}\hat{\boldsymbol{x}}_{(t/t-1)}) \\ \boldsymbol{P}_{(t)} = \boldsymbol{P}_{(t/t-1)} + \boldsymbol{K}_{(t)}\boldsymbol{H}_{(t)}\boldsymbol{P}_{(t/t-1)} \end{cases} \quad (6)$$

where $\boldsymbol{K}_{(t)} = \boldsymbol{P}_{(t/t-1)}\boldsymbol{H}_{(t)}{}^{T}\boldsymbol{S}_{(t/t-1)}{}^{-1}$, and $\boldsymbol{S}_{(t/t-1)} = \boldsymbol{H}_{(t)}\boldsymbol{P}_{(t/t-1)}\boldsymbol{H}_{(t)}{}^{T} + \boldsymbol{R}_{(t)}$. $\boldsymbol{R}$ is the covariance of the measurement noise Δz.

In our experiment, the covariances of the plant and measurement noises in (3) and (4) are set at $\boldsymbol{Q}$ = diag (1.0 m^2/s^4, 1.0 m^2/s^4) and $\boldsymbol{R}$ = diag (0.01 m^2, 0.01 m^2), respectively, through trial and error.

In this paper, a moving object is assumed to move at an almost constant velocity, and it is tracked using the usual Kalman filter. If it moves with various different motions, such as moving at a constant speed, going or stopping suddenly, or turning suddenly, the use of multi-model-based tracking, such as an interacting-multiple-model estimator, can improve the tracking performance [29] [30].

To extract W_{meas} and L_{meas} from the moving measurements, the heading information of the tracked object is needed. As shown in Fig. 4, we extract two feature lines (green lines) from the moving measurements using the split-and-merge method [31] and RANSAC [32] and determine the heading of the tracked object from the orientation of the feature lines. When the two feature lines cannot be extracted, we determine the heading from the velocity estimate of the object using arctan ($\dot{y}/\dot{x}$).

VI. Data Association

To track objects in crowded environments, we apply data association (i.e., one-to-one or one-to-many matching of tracked objects and moving measurements). In the hierarchical cooperative-tracking method, data association is performed by sensor nodes, whereas in the centralized method, it is performed by a central server.

As shown in Fig. 5, a validation region (black rectangle) is set around the predicted position (black circle) of a tracked object. The validation region is rectangular, and its length and width are 0.5 m longer than those of the object estimated at the previous time step (green dashed rectangle).

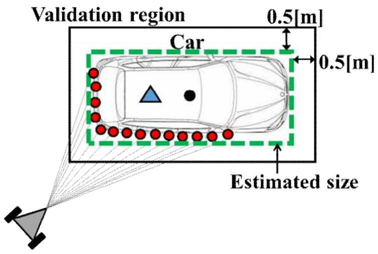


Figure 5. Moving measurements and data association.

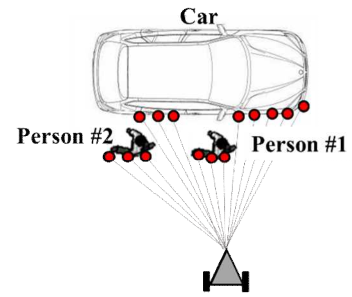
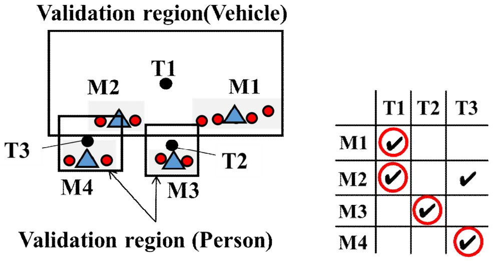


Figure 6. Moving measurements, in which two people move near a car.



	T1	T2	T3
M1	[✓]		
M2	[✓]		✓
M3		[✓]	
M4			[✓]

Figure 7. Data association for Fig. 6.

We refer to the representative point of clustered moving measurements (red circles) as the representative measurement (light blue triangle). The position of the representative measurement is the mean position of clustered moving measurements. Representative measurements inside the validation region are assumed to originate from the tracked object and are used to update the pose of the tracked object using (6), whereas those outside the validation region are identified as false alarms and discarded.

Figs. 6 and 7 illustrate an example of data association, in which two people move close to a car. In these figures, red circles indicate moving measurements, light blue triangles indicate representative measurements, and black circles indicate predicted positions of tracked objects. The right table in Fig. 7 shows the correspondence between tracked objects and representative measurements.

As shown in Fig. 7, multiple representative measurements are often obtained inside a validation region, and multiple validation regions also overlap. To achieve reliable data association (i.e., matching of tracked objects and representative measurements), we introduce the following rules:

1) Person: Because a person is small, he/she usually results in a representative measurement. Therefore, if a tracked object is assumed to be a person, one-to-one matching of a tracked person and a representative measurement is performed.

2) Vehicle (car, motorcycle, and bicycle): Because a vehicle is large, as shown in Fig. 6, it often results in several representative measurements. Thus, if a tracked object is assumed to be a vehicle, one-to-many matching of a tracked vehicle and representative measurements is performed.

As shown in Fig. 6, on urban streets, people often move close to vehicles, whereas vehicles move far away from each other. Thus, when representative measurements of people exist in the validation region of a tracked vehicle, they might be matched to the tracked vehicle. To avoid this situation, we begin data association with people.

We now detail our data association method from Fig. 7, in which the validation regions of a person (T3) and a car (T1) overlap. If tracked objects T2 and T3 are determined to be people from their estimated sizes (less than 0.8 m), the representative measurement M3 is matched with T2 and the representative measurement M4 nearest to T3 is matched with T3, both through one-to-one matching. Subsequently, if a tracked object T1 is determined to be a vehicle from the estimated size (larger than 0.8 m), the two representative measurements M1 and M2 in the validation region are matched with T1 through one-to-many matching. If validation regions of several people overlap, one-to-one matching is performed using the global nearest neighbor (GNN) method [12] [33].

Moving objects appear in and disappear from the sensing area of the sensor node. They are also occluded by each other and other objects in an environment. To maintain reliable tracking under such conditions, we implement following tracking rules.

1) Tracking initiation: If a representative measurement that is not matched with any tracked objects exists, it is assumed to either originate from a new object or to be an outlier. Therefore, we tentatively initiate tracking of the measurement with the Kalman filter. If the representative measurement remains visible in more than N_1 scans, it is assumed to originate from a new object and tracking is continued. If the representative measurements disappear within N_1 scans, it is assumed to be an outlier, and tentative tracking is terminated.

Because the size of the new tracked object is unknown at the initial time (scan), a rectangular validation region cannot be used for data association. Instead, we use a circular validation region with a constant radius of 2 m at the initial scan, and when the tracked object is matched with a representative measurement at the next scan, we estimate the size and decide whether the object is a vehicle or a person.

2) Tracking termination: When the tracked objects leave the sensing area of the sensor node or they meet occlusion, no representative measurements exist within their validation regions. If no measurements arise from the temporal occlusion, the measurements appear again. We thus predict the positions of the tracked objects using (5). If the representative measurements appear again within N_2 scans, we proceed with the tracking. Otherwise, we terminate the tracking.

In our experiments described in Section VIII, we set N_1 = 9 scans (0.9 s) and N_2 = 30 scans (3 s) through trial and error.

VII. Merging of Pose and Size Estimates by a Central Server

In the hierarchical cooperative-tracking method, each sensor node transmits the information of the tracked objects, which is shown in Table I (b), to the central server. When the central server receives this information from sensor nodes, it combines all the information together. It also merges the size, position and velocity of the moving objects locally estimated by the sensor nodes.

To combine the information, we apply data association (matching of tracking information). We present an example of our data association procedure in Figs. 8 and 9, in which two sensor nodes are tracking a car. In Fig. 8, red and blue rectangles indicate the sizes of the tracked objects #A (TA) and #B (TB), estimated by sensor nodes #1 and #2, respectively. Orange arrows indicate the headings of the objects.

If two tracked objects originate from the same object, the position, velocity and heading estimated by both sensor nodes will have similar values. Furthermore, if the tracked object is a vehicle, the size estimated by both sensor nodes will be large. If it is a person, the estimated size will be small. Therefore, we set a validation region with a constant radius of 3 m around the TA position (red star in Fig. 8) and match TB with TA by applying the following rules:

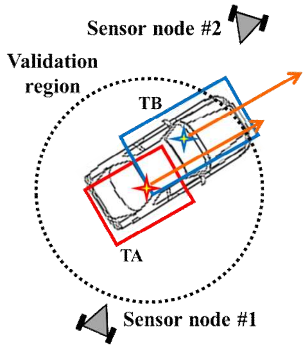


Figure 8. Data association of tracking information related to tracked objects TA and TB.

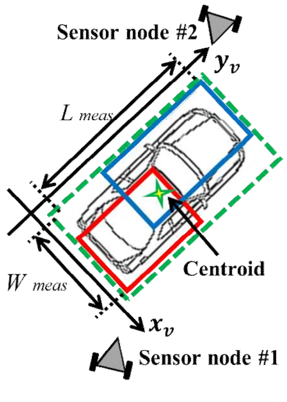
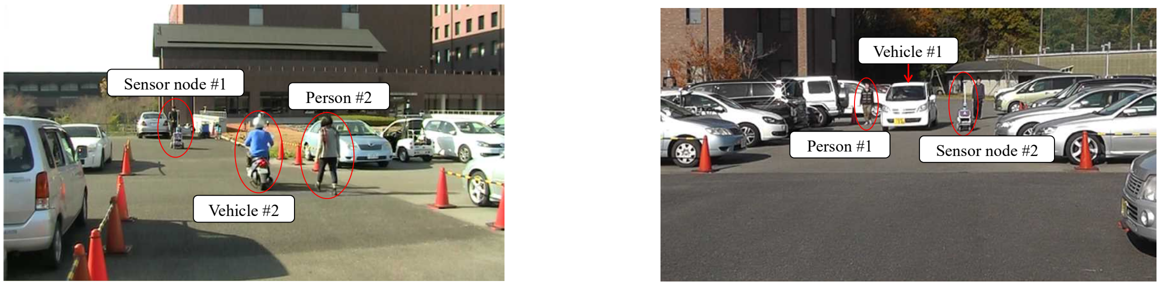


Figure 9. Merging of tracking information.



(a) Photo by camera #A

(b) Photo by camera #B

Figure 10. Photo of the experimental environment.

1) Same or different object: When the estimated position of TB (blue star) is located within the validation gate, and the differences in the velocity and heading estimates of TA and TB are less than D_1 m/s, and D_2°, respectively, the objects TA and TB are determined to originate from the same object. Otherwise, the objects TA and TB are determined to be different objects.

2) Vehicle or person: When the width and/or length estimates of the matched objects TA and TB are larger than 0.8 m, their objects are determined to originate from the same vehicle. When their width and length estimates are less than 0.8 m, the objects TA and TB are determined to originate from the same person.

When more than two tracked objects (e.g., TB and TC) are present in the validation region of TA, similar data association rules are applied. In our experiments described in Section VIII, we set D_1 = 0.8 m/s and D_2 = 15° through trial and error.

After the two tracked objects TA and TB have been matched, the tracking information is merged. As shown in Fig. 9, we select the tracked object TB, which has a larger rectangle (blue rectangle) than TA (red rectangle), and define an $x_v y_v$-coordinate frame on which the y_v-axis aligns with the heading of TB. Subsequently, a rectangle (the green dashed rectangle in Fig. 9) is then generated that encloses two rectangles of TA and TB using the position information of their vertices. We then estimate the size of the integrated object using (1) from the width and length of the new rectangle.

From the centroid position (green star) of the new rectangle, the position and velocity of the integrated object is estimated using the Kalman filter ((5) and (6)) under the assumption that the object is moving at an almost constant velocity.

VIII. Experimental Results and Discussion

A. Tracking by Two Mobile Sensor Nodes

We evaluated our cooperative-tracking method by conducting an experiment in a parking environment, as shown in Fig. 10. Two mobile sensor nodes tracked a car (vehicle #1), a motorcycle (vehicle #2), and two pedestrians (persons #1 and #2). Fig. 11 shows the movement paths of the sensor nodes (black dashed lines), vehicles #1 and #2 (blue and green lines), and persons #1 and #2 (red and black lines). The moving speeds of the sensor nodes, car, motorcycle, and persons were approximately 1.5, 15, 20, and 6 km/h, respectively.

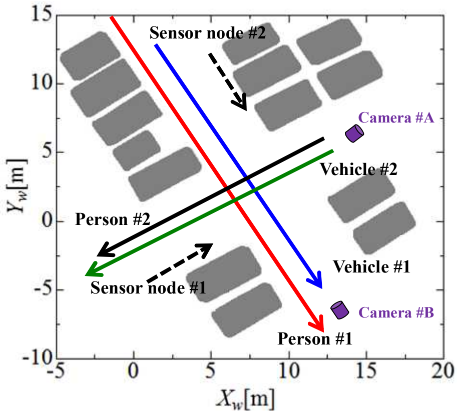
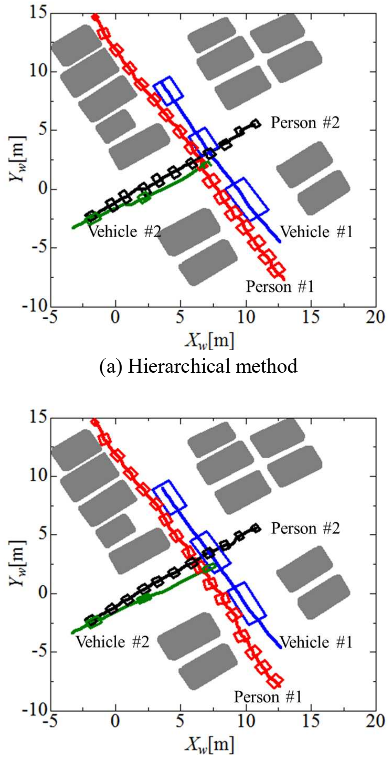


Figure 11. Movement paths of sensor nodes and moving objects.

Fig. 12 (a) shows the results of the position and size estimated using hierarchical cooperative tracking. We plot the estimated rectangles every 1 s (10 scans). Fig. 12 (b) shows the results of our previous centralized cooperative-tracking method. For comparison, individual tracking by each sensor node was also conducted. The tracking results for sensor nodes #1 and #2 are shown in Figs. 13 (a) and (b), respectively.

The estimated sizes of the car (vehicle #1) using cooperative and individual tracking are shown in Figs. 14 and 15, respectively. In these figures, red and blue lines indicate the estimated length and width, respectively. Two dashed lines indicate the true length and width of the car.

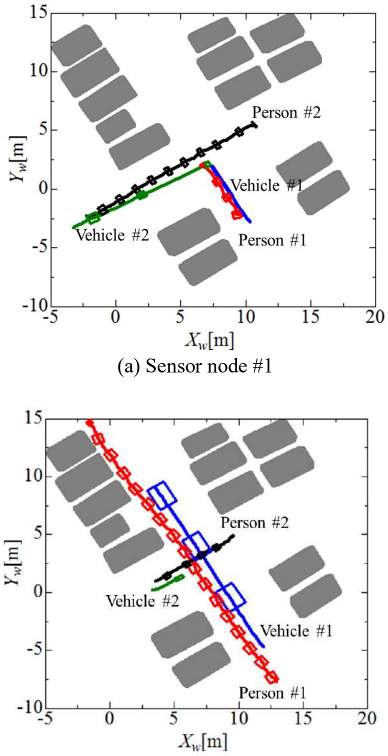
In individual tracking (Figs. 13 and 15), each sensor node partially tracks moving objects because the objects leave the sensing area of the sensor nodes and are blocked by parked cars. In contrast, in cooperative tracking, they always track their moving objects (Figs. 12 and 14) because the two sensor nodes share tracking data. It is clear from these figures that cooperative tracking provides better tracking accuracy than individual tracking.



(a) Hierarchical method

(b) Centralized method

Figure 12. Estimated track and size of moving objects using cooperative tracking.



(a) Sensor node #1

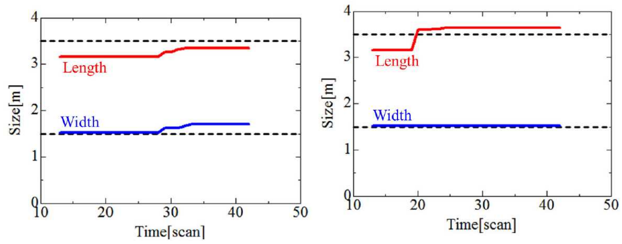
(b) Sensor node #2

Figure 13. Estimated track and size of moving objects using individual tracking.

As described in Section VI, when new moving objects appear in the sensing area of the sensor node, our tracker uses 9 scans (0.9 s) for the track initiation and begins to track the new objects from the 10th scan (1 s). In the 9 scans for the track initiation, vehicles #1 and #2 (car and motorcycle) have already moved over a long distance. This is the reason why the estimated tracks of vehicles #1 and #2 in Figs. 12 and 13 are shorter than their true movement paths shown in Fig. 11.

As shown in Fig. 14, the car (vehicle #1) size estimated using hierarchical and centralized cooperative-tracking methods are different. In the experiment, sensor node #2 detected vehicle #1 after 3 scans and began to track it from the 13th scan, whereas sensor node #1 detected vehicle #1 after 20 scans and began to track it from the 30th scan. In hierarchical cooperative tracking, each sensor node locally tracks the vehicle. The track initiation for vehicle #1 was executed in 3–13 scans by sensor node #2 and in 20–30 scans by sensor node #1. Therefore, the server received tracking information from sensor node #1 at the 30th scan and merged the information together.

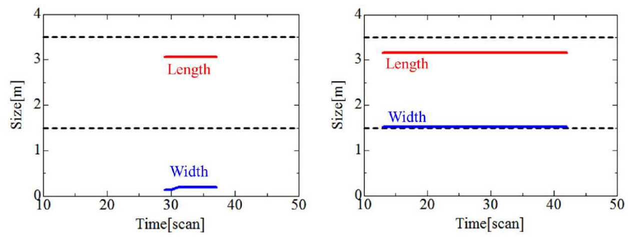
On the other hand, with centralized cooperative tracking, the central server estimates the size based on the moving measurements sent from the sensor nodes. Therefore, the track initiation for vehicle #1 was executed in only 3–13 scans by sensor node #2. When the server received the moving measurements from sensor node #1 at the 20th scan,



(a) Hierarchical method

(b) Centralized method

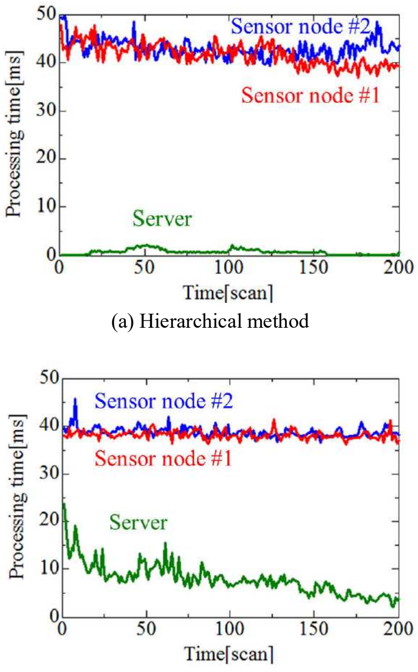
Figure 14. Estimated size of car (vehicle #1) using cooperative tracking.



(a) Sensor node #1

(b) Sensor node #2

Figure 15. Estimated size of car (vehicle #1) using individual tracking.



(a) Hierarchical method

(b) Centralized method

Figure 16. Processing time of sensor nodes and central server.

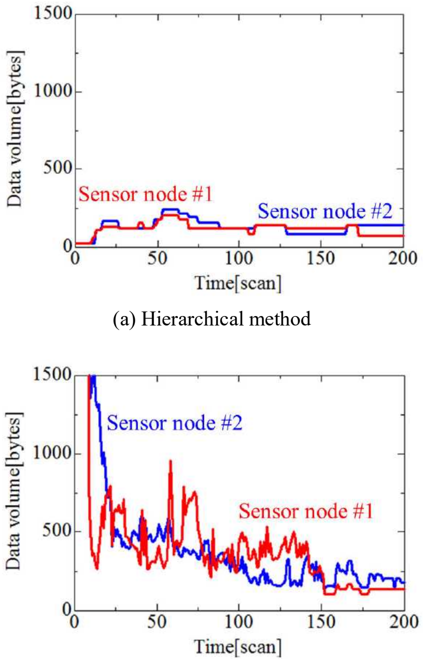
TABLE II. PROCESSING TIME OF SENSOR NODES AND CENTRAL SERVER

(a) HIERARCHICAL METHOD

	Max. [ms]	Min. [ms]	Mean [ms]
Central server	2.3	0.1	0.8
Sensor node #1	47.9	36.8	41.7
Sensor node #2	49.8	39.3	43.0

(b) CENTRALIZED METHOD

	Max. [ms]	Min. [ms]	Mean [ms]
Central server	23.8	2.2	7.9
Sensor node #1	41.5	36.1	38.2
Sensor node #2	45.7	36.5	38.8



(a) Hierarchical method

(b) Centralized method

Figure 17. Data volume sent to central server from sensor nodes.

TABLE III. DATA VOLUME SENT TO CENTRAL SERVER FROM SENSOR NODES

(a) HIERARCHICAL METHOD

	Max. [byte]	Min. [byte]	Mean [byte]
Sensor node #1	204	24	30
Sensor node #2	240	24	28

(b) CENTRALIZED METHOD

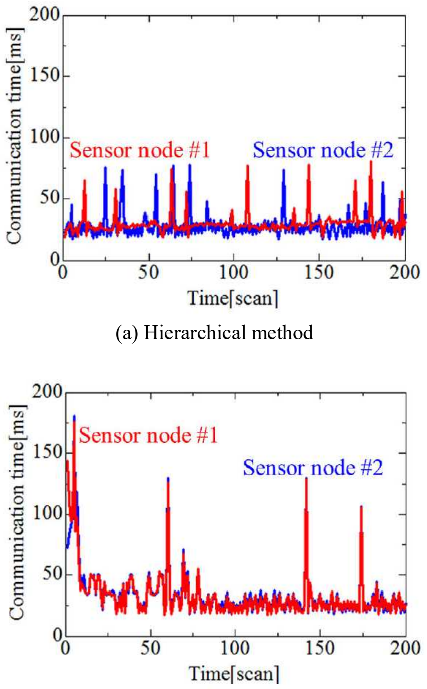
	Max. [byte]	Min. [byte]	Mean [byte]
Sensor node #1	3628	104	477
Sensor node #2	4952	136	540

it quickly merged the measurements with those from sensor node #2 and estimated the car size. As a result, compared to centralized cooperative tracking, hierarchical cooperative tracking causes a slight time lag when merging the information from both sensor nodes. This is why the car (vehicle #1) sizes estimated using hierarchical and centralized cooperative tracking are different.

In our experimental system, the model of the computers used in the sensor nodes and central server is Iiyama 15X7100-i7-VGB with a 2.8 GHz Intel core i7-4810MQ processor, and the operating system used is Microsoft Windows 7 Professional. We examined the processing time of the sensor nodes and the central server in the experiment.

The results for hierarchical and centralized cooperative tracking are shown in Fig.16 and Table II. In centralized cooperative tracking, the central server estimates the poses and sizes of moving objects based on the moving-object measurements sent from the sensor nodes. On the contrary, in hierarchical cooperative tracking, the sensor nodes estimate the poses and sizes of moving objects, and the central server merges their estimates. Therefore, compared with centralized cooperative tracking, hierarchical cooperative tracking reduces the computational burden on the central server.

Fig. 17 and Table III show the data volume sent to the central server from the sensor nodes in the experiment. Fig. 18 and Table IV also show the communication time required from the sensor nodes to the central server. In centralized cooperative tracking, sensor nodes upload the information shown in Table I (a), whereas in hierarchical cooperative tracking, sensor nodes upload the information shown in Table I (b), to the central server. It is clear from these figures and tables that the data volume and communication time for hierarchical cooperative tracking is less than that for centralized cooperative tracking.



(a) Hierarchical method

(b) Centralized method

Figure 18. Communication time required from sensor nodes to central server.

TABLE IV. COMMUNICATION TIME FROM SENSOR NODES TO CENTRAL SERVER

(a) HIERARCHICAL METHOD

	Max. [ms]	Min. [ms]	Mean [ms]
Sensor node #1	81.2	17.5	30.1
Sensor node #2	77.9	17.2	28.1

(b) CENTRALIZED METHOD

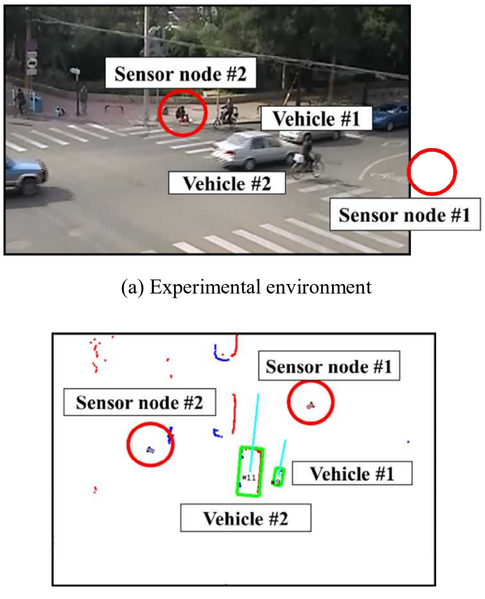
	Max. [ms]	Min. [ms]	Mean [ms]
Sensor node #1	175.9	17.7	33.8
Sensor node #2	180.3	18.0	34.0

B. *Tracking by Two Static Sensor Nodes*

We evaluated the accuracy of the pose and size estimates when using our cooperative-tracking method. For this purpose, we used Zhao's data set [34]. As shown in Fig. 19 (a), two laser scanners (SICK LMS200) were set at the height of 0.4 m in an intersection environment, and laser measurements were captured every 26 ms. We assumed that their measurements were captured by two sensor nodes and evaluated the tracking performance. The experimental duration was 108 s (4154 scans). Fig. 19 (b) shows the tracking result at 770 scans, where the green rectangles indicate the estimated size, and the light blue lines indicate the estimated heading. Red and blue dots indicate the laser measurements captured by sensors #1 and #2, respectively.

We examined the tracking performance for objects moving in the central area of the intersection (blue area in Fig. 20) where the sensing areas of the two sensor nodes overlapped.

Tables V and VI show the performance of the pose and size estimates when using cooperative and individual tracking, respectively. 'Actual objects' in tables were identified from camera images. 'Correct estimate of pose' means that the tracking method could always maintain a correct pose estimate of objects moving in the central area of the intersection, whereas 'incorrect estimate of pose' means that they failed in estimating the position. As described in Section V, the estimated size of the tracked object is used to classify the object as a person or a vehicle (i.e., car, motorcycle, and bicycle). If the estimated size in length or width is larger than 0.8 m, the object is determined to be a vehicle. If it is less than 0.8 m, the object is determined to be a person. In Tables V and VI, 'Correct



(a) Experimental environment

(b) Tracking result

Figure 19. Photo of the experimental environment and tracking result after 770 scans.

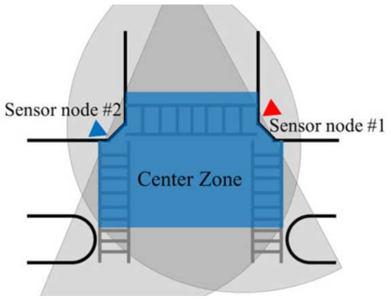


Figure 20. Overlapping sensing areas of two sensor nodes.

TABLE V. THE NUMBER OF OBJECTS WHOSE POSE (OR SIZE) ARE ESTIMATED CORRECTLY AND INCORRECTLY USING COOPERATIVE TRACKING

(a) HIERARCHICAL METHOD

		Correct estimate of pose (size)	Incorrect estimate of pose (size)
Actual object	Person	2 (2)	2 (0)
	Bicycle	20 (20)	3 (0)
	Car	30 (30)	7 (0)

(b) CENTRALIZED METHOD

		Correct estimate of pose (size)	Incorrect estimate of pose (size)
Actual object	Person	2 (2)	2 (0)
	Bicycle	23 (21)	0 (2)
	Car	31 (31)	6 (0)

TABLE VI. THE NUMBER OF OBJECTS WHOSE POSE (OR SIZE) ARE ESTIMATED CORRECTLY AND INCORRECTLY USING INDIVIDUAL TRACKING

(a) SENSOR NODE #1

		Correct estimate of pose (size)	Incorrect estimate of pose (size)
Actual object	Person	2 (2)	2 (0)
	Bicycle	18 (14)	5 (4)
	Car	26 (26)	11 (0)

(b) SENSOR NODE #2

		Correct estimate of pose (size)	Incorrect estimate of pose (size)
Actual object	Person	2 (2)	2 (0)
	Bicycle	18 (17)	5 (1)
	Car	22 (22)	15 (0)

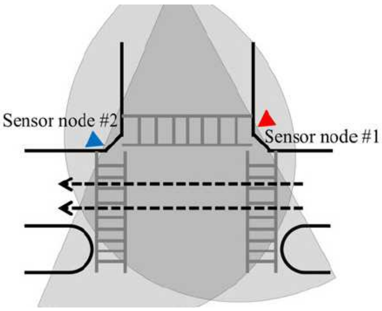


Figure 21. Movement paths of cars.

estimate of size' means that the tracking method could always maintain a correct classification of objects moving in the central area of the intersection, whereas 'Incorrect estimate of size' means that they failed in the classification.

It is clear from these tables that cooperative tracking provides better tracking accuracy than individual tracking. The performance of the pose and size estimates when using hierarchical cooperative tracking is slightly inferior to that of the centralized method. The pose estimation of cars using cooperative and individual tracking deteriorates when compared with people and bicycles. When cars moved along the paths shown by the dashed lines in Fig. 21, sensor nodes captured laser measurements related only to the right side of each car. In addition, they only partially captured laser measurements of cars due to occlusions. These cause incorrect pose estimation of cars.

IX. CONCLUSION AND FUTURE WORK

This paper presented a laser-based cooperative-tracking method for moving objects (vehicles and people) using multiple mobile sensor nodes located in the vicinity of the objects. In cooperative tracking, nearby sensor nodes share the tracking information. Hence, they enable the constant tracking of objects that may be invisible or partially visible to an individual sensor node.

We treated people and vehicles as rigid bodies and estimated both the poses and sizes of the objects. In a crowded environment, a vehicle can be occluded or rendered partially visible to each sensor node. To correctly estimate the vehicle's size, the laser measurements captured by sensor nodes have to be merged well. Therefore, we presented hierarchical cooperative-tracking method. Each sensor node obtained measurements related to the moving objects (moving measurements) and locally estimated the poses and sizes of the moving objects from the moving measurements using a Bayesian filter. It then sent these estimates to the central server, which then merged the pose and size estimates.

The performance of the hierarchical cooperative-tracking method was evaluated from two experimental results in outdoor environments by comparing it with a previous centralized method. The hierarchical method provided slightly inferior tracking accuracy than the centralized method. However, it had a smaller data volume sent to the central server from sensor nodes and a smaller computational cost in the central server than the centralized method. Therefore, the hierarchical method makes the tracking system scalable and robust.

In this paper, single-layer laser scanners were applied to sense the surrounding environment using mobile sensor nodes. Multilayer laser scanners can also capture height information from objects, and thus enable more accurate recognition of the surrounding environment than single-layer laser scanners. Current research is directed to the design of cooperative tracking by multiple sensor nodes equipped with multilayer laser scanners. To achieve cooperative tracking, the sensor nodes should always identify their own poses with a high degree of accuracy in a world coordinate frame. In this paper, we applied localization methods using dead reckoning and RTK-GPS. However, in city-canyon environments, the performance of localization using GPS deteriorates due to GPS multipath errors, diffraction problems and so on. To address this problem, we will embed a cooperative-localization method into our tracking system.

ACKNOWLEDGMENT

This study was partially supported by the Scientific Grants #26420213, the Japan Society for the Promotion of Science (JSPS) and the MEXT-Supported Program for the

Strategic Research Foundation at Private Universities, 2014–2018, Ministry of Education, Culture, Sports, Science and Technology, Japan.

We would also like to thank Professor Huijing Zhao of Peking University for providing the laser-scanning and vision data set for experiments in Section VII-B.

REFERENCES

[1] Y. Tamura, R Murabayashi, M. Hashimoto, and K. Takahashi, “Laser-based Cooperative Estimation of Pose and Size of Moving Objects using Multiple Mobile Robots,” Proc. of the Fifth Conf. on Intelligent Systems and Applications (INTELLI 2016), pp. 13–19, 2016.

[2] K. O. Arra and O. M. Mozos, Special issue on: People Detection and Tracking, Int. J. of Social Robotics, vol.2, no.1, pp. 1–107, 2010.

[3] C. Mertz, et al., “Moving Object Detection with Laser Scanners,” J. of Field Robotics, vol.30, pp. 17–43, 2013.

[4] T. Ogawa, H. Sakai, Y. Suzuki, K. Takagi, and K. Morikawa, “Pedestrian Detection and Tracking using In-vehicle Lidar for Automotive Application,” Proc. of IEEE Intelligent Vehicles Symp. (IV2011), pp. 734–739, 2011.

[5] A. Mukhtar, L. Xia, and T.B. Tang, “Vehicle Detection Techniques for Collision Avoidance Systems: A Review,” IEEE Trans. on Intelligent Transportation Systems, vol. 16, pp. 2318–2338, 2015.

[6] H. Cho, Y. W. Seo, B.V.K. V. Kumar, and R. R. Rajkumar, “A Multi-sensor Fusion System for Moving Object Detection and Tracking in Urban Driving Environments,” Proc. of Int. Conf. on IEEE Robotics and Automation (ICRA2014), pp. 1836–1843, 2014.

[7] D. Z. Wang, I. Posner, and P. Newman, “Model-free Detection and Tracking of Dynamic Objects with 2D Lidar,” Int. J. of Robotics Research, vol.34, pp. 1039–1063, 2015.

[8] D. Z. Wang, I. Posner, P. Newman, “What could move? Finding cars, pedestrians and bicyclists in 3D laser data,” Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA2012), pp. 4038–4044, 2012.

[9] Z. Yan, N. Jouandeau, and A. A. Cherif, “A Survey and Analysis of Multi-Robot Coordination,” Int. J. of Advanced Robotic Systems, vol. 10, pp. 1–18, 2013.

[10] S. Nadarajah and K. Sundaraj, “A Survey on Team Strategies in Robot Soccer: Team Strategies and Role Description,” Artificial Intelligence Review, vol. 40, pp. 271–304, 2013.

[11] K. Kakinuma, M. Hashimoto, and K. Takahashi, “Outdoor Pedestrian Tracking by Multiple Mobile Robots based on SLAM and GPS Fusion,” Proc. of IEEE/SICE Int. Symp. on System Integration (SII2012), pp. 422–427, 2012.

[12] M. Ozaki, K. Kakinuma, M. Hashimoto, and K. Takahashi, “Laser-based Pedestrian Tracking in Outdoor Environments by Multiple Mobile Robots,” Sensors, vol. 12, pp. 14489–14507, 2012.

[13] S.J. Julier and J.K. Uhlmann, “A Non-divergent Estimation Algorithm in the Presence of Unknown Correlations,” Proc. of the IEEE American Control Conf., pp. 2369–2373, 1997.

[14] F. Fayad and V. Cherfaoui, “Tracking Objects using a Laser Scanner in Driving Situation based on Modeling Target Shape,” Proc. of the 2007 IEEE Int. Vehicles Symp. (IV2007), pp. 44–49, 2007.

[15] T. Miyata, Y. Ohama, and Y. Ninomiya, “Ego-Motion Estimation and Moving Object Tracking using Multi-layer LIDAR,” Proc. of IEEE Intelligent Vehicles Symp. (IV2009), pp. 151–156, 2009.

[16] K. Granstrom, C. Lundquist, F. Gustafsson, and U. Orguner, “Radom Set Methods, Estimation of Multiple Extended Objects,” IEEE Robotics & Automation Magazine, pp. 73–82, June 2014.

[17] L. Mihaylova, et al., “Overview of Bayesian Sequential Monte Carlo Methods for Group and Extended Object Tracking,” Digital Signal Processing, vol. 25, pp.1–16, 2014.

[18] J. Lan and X. R. Li, “Tracking of Extended Object or Target Group using Random Matrix Part I: New Model and Approach,” Proc. of 15th Int. Conf. on Information Fusion (FUSION2012), pp.2177–2184, 2012.

[19] Z.Wang and D. Gu, “Cooperative Target Tracking Control of Multiple Robots,” IEEE Trans. on Industrial Electronics, vol. 59, pp. 3232–3240, 2012.

[20] K. Zhou and S. I. Roumeliotis, “Multirobot Active Target Tracking with Combinations of Relative Observations,” IEEE Trans. on Robotics, vol. 27, pp. 678–695, 2011.

[21] A. Ahmad and P. Lima, “Multi-robot Cooperative Spherical-Object Tracking in 3D Space based on Particle Filters,” Robotics and Autonomous Systems, vol. 61, pp. 1084–1093, 2013.

[22] P. U. Limaa, et al., “Formation Control Driven by Cooperative Object Tracking,” Robotics and Autonomous Systems, vol. 63, Part 1, pp. 68–79, 2015.

[23] C. Robin and S. Lacroix, “Multi-robot Target Detection and Tracking: Taxonomy and Survey,” Autonomous Robots, vol. 40, pp. 729–760, 2016.

[24] C. T. Chou, J. Y. Li, M. F. Chang, and L. C. Fu, “Multi-Robot Cooperation Based Human Tracking System Using Laser Range Finder,” Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA2011), pp. 532–537, 2011.

[25] N. A. Tsokas and K. J. Kyriakopoulos, “Multi-robot Multiple Hypothesis Tracking for Pedestrian Tracking,” Autonomous Robot, vol. 32, pp. 63–79, 2012.

[26] M. Hashimoto, R. Izumi, Y. Tamura, and K. Takahashi, “Laser-based Tracking of People and Vehicles by Multiple Mobile Robots,” Proc. of the 11th Int. Conf. on Informatics in Control, Automation and Robotics (ICIT2014), pp. 522–527, 2014.

[27] M. Hashimoto, S. Ogata, F. Oba, and T. Murayama, “A Laser-based Multi-Target Tracking for Mobile Robot,” Intelligent Autonomous Systems 9, pp. 135–144, 2006.

[28] Y. Bar-Shalom and T. E. Fortmann, “Tracking and Data Association,” Academic Press,Inc., 1988.

[29] H.A.P.Blom and Y.Bar-Shalom, “The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficient,” IEEE Trans. on Automatic Control, vol.33, pp.780–783, 1988.

[30] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, “Interacting Multiple Model Methods in Target Tracking: A Survey,” IEEE Trans. on Aerospace and Electronic Systems, vol.34, pp.103–123, 1998.

[31] V. Nguyen, A. Martinelli, N. Tomatis, and R. Siegwart, “A Comparison of Line Extraction Algorithms using 2D Laser Rangefinder for Indoor Mobile Robotics,” Proc. of 2005 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS2009), pp. 1929–1934, 2009.

[32] M. Fischler and R. Bolles, “Random Sample Consensus: A Paradigm for Model Fitting Applications to Image Analysis and Automated Cartography,” Proc. of Image Understanding workshop, pp. 71–88, 1980.

[33] P. Konstantinova, A. Udvarev, and T. Semerdjiev, “A Study of a Target Tracking Algorithm Using Global Nearest Neighbor Approach,” Proc. of Int. Conf. on Systems and Technologies, 2003.

[34] H. Zhao, “Open Resource, Networked horizontal laser scan data at intersection (20071012),” available from <http://www.poss.pku.edu.cn/download.html>, (accessed on 12 December 2014).