LiDAR-Based Cooperative Scan Matching for Relative Pose Estimation of Multiple Vehicles in GNSS-Denied Environments

Ryoga Takahashi
Graduate School of Science and Engineering
Doshisha University
Kyotanabe, Kyoto 610-0394 Japan

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 relative pose estimation method for multiple vehicles that involves scan matching using Light Detection And Ranging (LiDAR) measurements captured from overlapping sensing areas of neighboring vehicles. Such Cooperative Scan Matching (CSM) can be used to Cooperative Moving-Object Tracking (CMOT) by vehicles in Global Navigation Satellite Systems (GNSS)-denied environments. Each vehicle is equipped with a LiDAR and detects static measurements originating from static objects, such as building walls and utility poles, from its own LiDAR scan images by applying an occupancy grid method. Each vehicle then obtains point feature histograms related to static features using Fast Point Feature Histograms (FPFH). The information related to LiDAR measurements and point feature histograms is sent to a central server. After the central server matches the environmental features obtained by the vehicles based on point feature histograms and estimates the relative pose of the vehicles by RAaNdom Sample Consensus (RANSAC) based coarse registration and Normal Distributions Transform (NDT) scan matching-based fine registration. Experimental results using two LiDAR-equipped vehicles in two road environments show that the proposed CSM has better applicability in urban environments with a high number of streets.**

Keywords-multi-vehicles; LiDAR; cooperative scan matching; relative pose estimation; FPFH; NDT scan mathcing.

I. INTRODUCTION

Tracking (i.e., estimating the position, size, and velocity) of multiple moving objects such as people, cars, bicycles, and motorcycles in real environments is an important issue for the safe navigation and autonomous driving of mobile robots and vehicles [1]–[3]. For such Moving-Object Tracking (MOT), the use of stereo camera, Light Detection And Ranging (LiDAR), and radar in mobile robotics and Intelligent Transportation Systems (ITS) has been drawing considerable interest. This paper focuses on LiDAR-based MOT.

MOT using only a LiDAR mounted on a vehicle (called Individual MOT (IMOT)) cannot track moving objects that exist outside the sensing area of the vehicle or in a blind spot of the vehicle. To address this problem of IMOT, Cooperative Moving-Object Tracking (CMOT) [4] and cooperative perception (CP) [5][6] have been presented. Multiple vehicles exchange LiDAR sensing data through a vehicle-to-vehicle communication network, and they detect and track moving objects in each other's blind spots. CMOT and CP can then improve the accuracy and reliability of MOT and environmental sensing.

In MOT, the occupancy grid method [7] is usually used to detect moving objects from the LiDAR measurements. In this method, the accurate self-pose (position and attitude) of the LiDAR-equipped vehicle in the world coordinate frame is required to map the LiDAR measurements acquired in the sensor coordinate frame onto the grid map in the world coordinate frame. In CMOT, if the accuracy of self-pose estimation in the world coordinate frame differs from vehicle to vehicle, each vehicle would misrecognize the same object as two different objects. In addition, it would recognize an object's size as larger or smaller than the actual size.

For accurate self-pose estimation of vehicles in CMOT, we used Real Time Kinematic Global Navigation Satellite Systems (RTK GNSS) under the assumption that CMOT is applied in outdoor open-sky environments [8]. However, GNSS, including RTK GNSS, lack credibility due to the multipath or mask effects in urban street canyons. CMOT can be performed in such GNSS-denied environments using Cooperative Scan Matching (CSM) [9] and cooperative positioning [10][11]. CSM estimates the relative pose of vehicles by matching the LiDAR measurements in the overlapping sensing areas of these vehicles, and the self-poses of the vehicles are corrected using the relative pose estimate. Iterative Closest Point (ICP) [12] or Normal Distributions Transforms (NDT) -based scan matching [13] are used as CSM algorithms.

In relative pose estimation using ICP or NDT-based scan matching (called fine registration), an appropriate initial value of the relative pose (called coarse registration) should be calculated to avoid the local-minimum problem in the iterative calculation in the fine registration. Our previous work [9] extracted pole-like objects in environments, such as utility poles and light poles, from the LiDAR measurements and utilized them as environmental features for coarse registration. However, many environments do not have such objects. In addition, where they do exist, they are frequently occluded by surrounding moving objects, such as cars, tracks, and buses.

This paper presents a coarse registration method that uses Fast Point Feature Histograms (FPFH) and Random Sample Consensus (RANSAC)-based algorithm [14][15]. Since such coarse registration does not limit the environmental features to pole-like objects, it can be applied in various GNSS-

 ISBN: 978-1-61208-917-1

denied environments, such as urban city roads and roads with many street trees. In this paper, NDT scan matching-based fine registration is used to estimate the relative pose of vehicles. In addition, a success/failure decision method of CSM is presented.

The rest of this paper is organized as follows: Section II provides an overview of our experimental system and CMOT method. Section III describes the relative pose estimation by CSM and the success/failure decision. In Section IV, experimental results obtained in outdoor road environments by two vehicles (small electric car and motorcycle) are shown, followed by our conclusions in Section V.

II. EXPERIMENTAL SYSTEM AND CMOT OVERVIEW

In this section, an overview of our experimental system and CMOT is presented.

A. Experimental System

Figure 1 shows the two experimental vehicles: a small electric car (Toyota auto body, COMS) and a motorcycle (Honda, Gyro Canopy). A 32-layer scanning LiDAR (Velodyne, HDL-32E) is installed on the top of each vehicle. The maximum range of the LiDAR is 70 m, the horizontal viewing angle is 360° with a resolution of 0.16°, and the vertical viewing angle is 41.34° with a resolution of 1.33°. The LiDAR provides 384 measurements (the object's 3D position and reflection intensity) every 0.55 ms (at 2° horizontal angle increments). The period needed for the laser beam to complete one rotation (360°) in the horizontal direction is 100 ms; thus 70,000 measurements are obtained in one rotation.

The motorcycle (vehicle 2) is also equipped with Inertial Measurement Unit (IMU) (Xsens, MTi-300). The IMU outputs the attitude angle (roll and pitch angles) and angular velocity (roll, pitch, and yaw angular velocities) every 10 ms. The errors in attitude angle and angular velocity are less than ±0.3° and ±0.2°/s, respectively.

B. Overview of CMOT

Figure 2 shows a sequence of CMOT. Each vehicle senses its surrounding environment with its scanning LiDAR. The LiDAR obtains range measurements by scanning laser beams. Thus, when the vehicle moves, the entire scan data within one scan (laser beam rotation of 360° in the horizontal plane) cannot be obtained at the same pose of the vehicle. Therefore, if the entire scan data obtained within one scan are mapped onto the world coordinate frame using the self-pose information of the vehicle, distortion will arise in the LiDAR scan images. To correct this distortion, the vehicle's pose is determined in a period shorter than than the LiDAR scan period, i.e., for every LiDAR measurement (0.55 ms) in the scan. Vehicle 1 (small car) corrects the distortion in the LiDAR scan images using the information from only LiDAR measurements [16]. Vehicle 2 (motorcycle) corrects the distortion using the information from its LiDAR and IMU measurements because it changes its pose more significantly compared with vehicle 1 [17].

(a) Vehicle 1 (small car)

(b) Vehicle 2 (scooter)

Figure 1. Overview of the experimental vehicle.

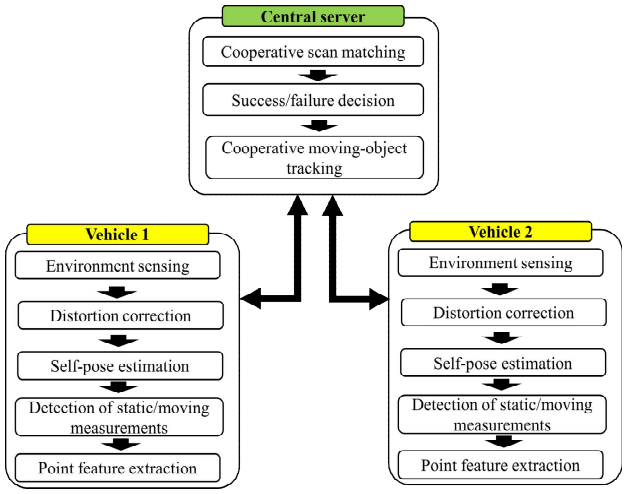


Figure 2. Overview of the cooperative moving-object tracking (CMOT).

Each vehicle maps their corrected LiDAR measurements onto the grid map. Based on the occupancy grid method, the LiDAR measurements are classified into two types, namely moving and static measurements, which originate from moving and static objects, respectively. Point feature histograms are obtained using FPFH for static measurements.

 ISBN: 978-1-61208-917-1

The information of the LiDAR measurements is then uploaded to the central server. This information includes time stamps, 3D positions of static measurements with the point feature histograms, 3D positions of the moving measurements, and self-poses of the vehicles.

After receiving the information of the LiDAR measurements, the central server estimates the size, position, and velocity of the moving objects using a Bayesian-based algorithm [10]. The estimated information is then fed back to the vehicles.

In this paper, NDT-based Simultaneous Localization and Mapping (SLAM) [19] is assumed to estimate the self-poses of the vehicles. In CMOT, if two vehicles have different accuracies of their self-poses in the world coordinate frame, the mapping of LiDAR measurements for the same object onto the grid map will cause errors, and the CMOT performance will deteriorate. Therefore, the relative pose between two vehicles is estimated by CSM, and the self-poses of the vehicles are corrected using the relative pose information. In this paper, for simplicity, only the self-pose of vehicle 2 is corrected.

CSM estimates the relative pose between two vehicles by matching the static measurements in the overlapping sensing areas of the vehicles using NDT scan matching. However, if the initial pose given to NDT scan matching is inaccurate, NDT scan matching will output an inaccurate relative pose due to the local-minimum problem in its iterative calculation. Therefore, the static measurements in the overlapping sensing areas of two vehicles are matched based on the point feature histograms, and the relative pose is coarsely estimated using RANSAC. The relative pose estimate obtained by such coarse registration is given to NDT scan matching as the initial pose.

Finally, the success or failure of CSM is assessed. If it is deemed successful, the moving measurements detected by two vehicles are re-mapped to the grid map, and the moving objects are tracked.

III. COOPERATIVE SCAN MATCHING

In this section, the point feature histograms obtained using FPFH, which is a metric for matching static measurements from two LiDARs, are first described. Next, the relative pose between two vehicles is estimated using coarse and fine registration. Finally, the success/failure decision method of CSM is described.

A. Point Feature Histograms

Each vehicle obtains point features using FPFH [14] from the static measurements. Vehicle 1 first maps the static measurements onto a voxel map (grid size of 1 m) in its own vehicle coordinate frame Σ_1 (fixed to vehicle 1) and downsamples the static measurements using a voxel grid filter. The centroid of the static measurements in the i-th voxel (i = 1, 2, …) on the voxel map is then obtained. The centroid is called the feature point A_i. Vehicle 2 obtains the feature point B_i in the same way from the static measurements in its own coordinate frame Σ_2.

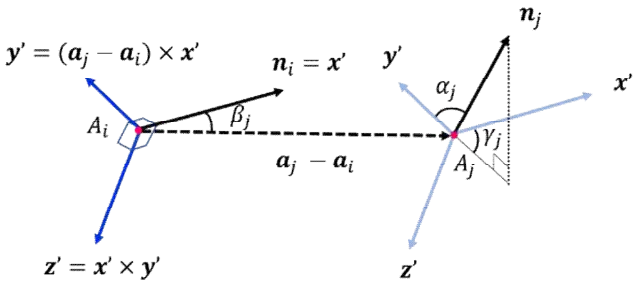


Figure 3. Point feature.

The point feature histograms are calculated based on the feature points A_i and B_i. Since it is the same for both vehicles, only the point feature histograms calculation method for vehicle 1 is described in this paper.

The 3D position of the feature point A_i in the i-th voxel is denoted by $\boldsymbol{a}_i$ in Σ_1. The 3D position of the feature point A_j in the j-voxel (j =1, 2, ... 124), which is located around the i-th voxel, is also denoted by $\boldsymbol{a}_j$ in Σ_1. As shown in Figure 3, the normal vectors for the feature points A_i and A_j are denoted by $\boldsymbol{n}_i$ and $\boldsymbol{n}_j$, respectively. A coordinate frame Σ_{Ai} is defined, in which the feature point A_i is used as the origin O_{Ai}, the normal vector $\boldsymbol{n}_i$ as the x' axis, $(\boldsymbol{a}_j - \boldsymbol{a}_i) \times \boldsymbol{x}'$ as the y' axis, and the axis orthogonal to the x' and y' axes as the z' axis.

Then, the triple feature (α_j, β_j, γ_j) related to the angles in Σ_{Ai} is defined by $\alpha_j = \boldsymbol{y}' \cdot \boldsymbol{n}_j$, $\beta_j = \boldsymbol{x}' \cdot (\boldsymbol{a}_j - \boldsymbol{a}_i)$, and $\gamma_j = \arctan(\boldsymbol{z}' \cdot \boldsymbol{n}_j / \boldsymbol{x}' \cdot \boldsymbol{n}_j)$. The point features $\boldsymbol{SPFH}(A_i)$ of 3×124-dimensional vector is obtained by calculating the features for the 124 feature points A_j (j =1, 2, ... 124) around A_i. In similar way, the point features $\boldsymbol{SPFH}(A_j)$ for the feature points A_j (j =1, 2, ... 124) around A_i are obtained. The final histograms (33 dimensions) related to the feature point A_i are then determined from

$$\boldsymbol{FPFH}(A_i) = \boldsymbol{SPFH}(A_i) + \frac{1}{124} \sum_{j=1}^{124} \frac{1}{w_j} \boldsymbol{SPFH}(A_j) \quad (1)$$

where the weight w_j is equal to $|\boldsymbol{a}_j - \boldsymbol{a}_i|$.

B. Relative Pose Estimation

In the coarse registration, the relative pose between vehicles is determined by matching the feature points A_i and B_i with similar point feature histograms. However, since many feature points have similar feature histograms, incorrect matching often occurs. Therefore, the following RANSAC-based method is used to perform the coarse registration.

Step 1: Three feature points A_i^* (i = 1, 2, 3) are randomly extracted from the set of feature points obtained by vehicle 1. Then, 100 feature points B_i (i = 1, … 100) with similar feature histograms as those of A_i^* are extracted using the k-nearest neighbor method from the set of feature points obtained by vehicle 2. Then, one feature point B_i^* is randomly selected from the 100 feature points B_i.

Step 2: The pose of vehicle 2 relative to vehicle 1 is denoted by $\delta \boldsymbol{X} = (\delta x, \delta y, \delta z, \delta\phi, \delta\theta, \delta\psi)^T$, where $\delta\boldsymbol{x}=(\delta x,\delta y,\delta z)^T$ and $\delta\boldsymbol{\theta}=(\delta\phi,\delta\theta,\delta\psi)^T$ are the relative position and attitude angle (roll, pitch, and yaw angles), respectively.

The centroid positions of three feature points (A_1^*, A_2^*, A_3^*) and (B_1^*, B_2^*, B_3^*) are denoted by $\bar{\boldsymbol{a}}$ and $\bar{\boldsymbol{b}}$, respectively. The feature point matrices are denoted by $\delta\boldsymbol{a}=(\delta\boldsymbol{a}_1,\delta\boldsymbol{a}_2,\delta\boldsymbol{a}_3)^T$ and $\delta\boldsymbol{b}=(\delta\boldsymbol{b}_1,\delta\boldsymbol{b}_2,\delta\boldsymbol{b}_3)^T$, where $\delta\boldsymbol{a}_i \triangleq \boldsymbol{a}_i^* - \bar{\boldsymbol{a}}$ and $\delta\boldsymbol{b}_i \triangleq \boldsymbol{b}_i^* - \bar{\boldsymbol{b}}$, and $\boldsymbol{a}_i^*$ and $\boldsymbol{b}_i^*$ are the 3D positions of the feature points A_i^* and B_i^*, respectively. With use of the matrices $\boldsymbol{U}$ and $\boldsymbol{V}$, which are defined by the singular value decomposition ($\boldsymbol{H} = \boldsymbol{U\Sigma V}^T$) of the matrix $\boldsymbol{H} = \delta\boldsymbol{b}\bullet\delta\boldsymbol{a}^T$, the relative position $\delta\boldsymbol{x}$ and the rotational matrix $\boldsymbol{R}(\delta\boldsymbol{\theta})$ related to the relative attitude angle $\delta\boldsymbol{\theta}$ are given by

$$\boldsymbol{R}(\delta\boldsymbol{\theta}) = \boldsymbol{V}\boldsymbol{U}^T = \begin{pmatrix} \cos\delta\theta\cos\delta\psi & \sin\delta\phi\sin\delta\theta\cos\delta\psi - \cos\delta\phi\sin\delta\psi & \cos\delta\phi\sin\delta\theta\cos\delta\psi + \sin\delta\phi\sin\delta\psi \\ \cos\delta\theta\sin\delta\psi & \sin\delta\phi\sin\delta\theta\sin\delta\psi + \cos\delta\phi\cos\delta\psi & \cos\delta\phi\sin\delta\theta\sin\delta\psi - \sin\delta\phi\cos\delta\psi \\ -\sin\delta\theta & \sin\delta\phi\cos\delta\theta & \cos\delta\phi\cos\delta\theta \end{pmatrix} \quad (2)$$

$$\delta\boldsymbol{x} = \bar{\boldsymbol{a}} - \boldsymbol{R}(\delta\boldsymbol{\theta})\bar{\boldsymbol{b}} \quad (3)$$

With use of the relative pose, the 3D position $\boldsymbol{b}_i$ of the feature point B_i in Σ_2 (obtained by vehicle 2) can be transformed to the 3D position $\boldsymbol{b}_i^{(1)}$ in Σ_1. The feature point nearest to $\boldsymbol{b}_i^{(1)}$ is extracted from the set of feature points A_i (i =1, 2, ...), and the 3D position of the nearest feature point in Σ_1 is denoted by $\tilde{\boldsymbol{a}}_i$. Then, the cost function is given by

$$J = \frac{1}{N_B}\sum_{i=1}^{N_B} (\tilde{\boldsymbol{a}}_i - \boldsymbol{b}_i^{(1)})^T (\tilde{\boldsymbol{a}}_i - \boldsymbol{b}_i^{(1)}) \quad (4)$$

where $\boldsymbol{b}_i^{(1)} = \boldsymbol{R}(\delta\boldsymbol{\theta})\boldsymbol{b}_i + \delta\boldsymbol{x}$. N_B is the number of the feature points B_i.

Step 3: Steps 1 and 2 are repeated 10,000 times to find the relative pose $\delta\boldsymbol{X}$ with the smallest J. Then, the relative pose $\delta\boldsymbol{X}_0$ is obtained in the coarse registration.

In NDT scan matching, the relative pose $\delta\boldsymbol{X}_0$ is used as the initial value, and the iterative calculation is performed using the Newton method to maximize the likelihood function in Eq. (5). The relative pose $\delta\boldsymbol{X}^*$ can then be obtained.

$$\Lambda = \prod_{j=1}^{N_2} \exp\left(-\frac{1}{2} (\boldsymbol{q}_j^{(1)} - \bar{\boldsymbol{p}}_i)^T \boldsymbol{\Omega}_i^{-1} (\boldsymbol{q}_j^{(1)} - \bar{\boldsymbol{p}}_i) \right) \quad (5)$$

where $\bar{\boldsymbol{p}}_i$ and $\boldsymbol{\Omega}_i$ are the mean and covariance, respectively, of the 3D positions of the static measurements in the i-th voxel of the voxel map in Σ_1. $\boldsymbol{q}_j$ is the 3D position of the j-th static measurement (j = 1, 2, $\cdots N_2$) obtained by vehicle 2 in Σ_2 and transformed to Σ_1.

C. Success/Failure Decision

Determination of CSM success or failure is vital (i.e. whether the relative pose obtained by CSM is correct or not).

After the NDT scan matching-based fine registration, the static measurements are first downsampled with a voxel grid filter (cell size of 0.5 m in this paper). The sets of downsampled static measurements related to vehicles 1 and 2 are called static measurement sets 1 and 2, respectively. For each measurement in static measurement sets 2, the closest measurement is identified from static measurement sets 1, and the Euclidean distance between the two measurements is calculated. This process is conducted M_2 times, where M_2 is the number of measurements in static measurement sets 2. The number M_{near} of static measurements whose Euclidean distance is 0.5 m or less is counted.

The matching rate is defined by $M_{near} \diagup M_2$. If the matching rate is equal to or greater than a threshold (33% in this paper), then the CSM is deemed successful, and the relative pose obtained by the CSM is utilized in CMOT.

The computational cost related to the coarse registration is high. Therefore, if the CSM is found to succeed in the current scan, the coarse registration will not be executed in the next scan. The relative pose estimated by the NDT scan matching in the current scan is then used as the initial pose for the NDT scan matching in the next scan. By contrast, if the CSM is found to have failed in the current scan, both the coarse and fine registration is utilized to estimate the relative pose in the next scan.

IV. Experimental Results

Experiments are conducted in two environments: an urban road environment with many buildings (environment 1; Figure 4 (a)) and a road environment with street trees (environment 2; Figure 5 (a)). In environment 1, vehicle 1 follows vehicle 2, which moves at about 40 km/h. The distance between the two vehicles is shown in Figure 4 (b). On the road, there are 20 cars and 5 pedestrians. In environment 2, vehicle 1 follows vehicle 2, which moves at about 30 km/h, and the distance between the two vehicles is shown in Figure 5 (b). On the road, there are 3 cars and 10 pedestrians.

Figure 6 shows the mapping results of the LiDAR measurements captured by the two vehicles at 40 s in environment 1. The mapping performance is compared in the following cases.

Case 1: Mapping using relative pose estimate by CSM (proposed method)
Case 2: Mapping using self-pose by RTK GNSS
Case 3: Mapping using self-pose by standard GNSS

The two experimental vehicles are equipped with RTK GNSS (Novatel ProPak-V3 on vehicle 1 and Novatel PwrPak7-E1 on vehicle 2) to evaluate our studies. In case 2, the LiDAR measurements captured by the two vehicles are mapped onto the world coordinate frame using the self-poses of the vehicles obtained by RTK GNSS. The GNSS in the normal mode (standard GNSS) is widely used in real-world

applications. However, the positioning accuracy of standard GNSS is worse than that of RTK GNSS. Thus, Gaussian errors with a mean of zero and standard deviation of 1 m are added to the self-pose obtained by RTK GNSS to generate the position information of standard GNSS. In case 3, LiDAR measurements are mapped onto the world coordinate frame using such pseudo-self-poses of the vehicles.

As shown in Figure 6 (d), in case 3, since the relative pose of the two vehicles is inaccurate, the LiDAR measurements related to the building walls and vehicles are not matched significantly. Then, the two detected vehicles will be recognized as four vehicles by our CMOT. In case 2 (Figure 6 (c)), although RTK GNSS gives an accurate relative pose between the two vehicles, the LiDAR measurements related to the building walls and vehicles are slightly mismatched. As shown in Figure 6 (b), the mapping performance of the proposed method (case 1) is better than that of the two other cases.

Figure 7 shows the mapping results of the LiDAR measurements captured by vehicles 1 and 2 at 14 s in environment 2. It is clear from this figure that the mapping performance of the proposed method (case 1) is better than that of the two other cases.

Table 1 shows the success rate of CSM and performance of CSM success/failure decision by the matching rate. Four measures are used to evaluate the performance of CSM success/failure decision: accuracy, precision, recall, and F-measure. The accuracy represents the ratio of the number of scans where the success/failure decision by the matching rate matches the actual success/failure result of CSM. Precision is the ratio of the number of scans where the CSM actually succeeds to the number of scans where the CSM is considered successful in the success/failure decision. Recall is the ratio of the number of scans where the CSM is found successful in the success/failure decision to the number of scans where the CSM actually succeeds. F-measure is the harmonic mean of precision and recall.

The success rate of CSM in environment 2 is lower than that in environment 1. As shown in Figures 4 (b) and 5 (b), the distance between the two vehicles in environment 2 is longer than that in environment 1. As the distance between the two vehicles increases, the overlapping area of the LiDARs decreases, and the inaccuracy of feature point matching increases. In addition, since the road environment with street trees (environment 2) has many similar point feature histograms, accurate feature point matching is difficult. Consequently, the success rate of CSM in environment 2 is lower than that in environment 1.

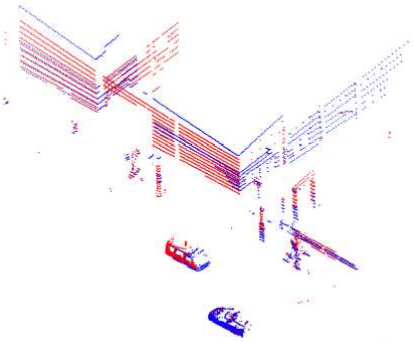
(a) Photo of environment (b) Distance between vehicles

Figure 4. Photo of environment and distance between vehicles (environment 1). The left photo is the bird's-eye view, and the yellow line in the left photo indicates the movement path of the vehicles.

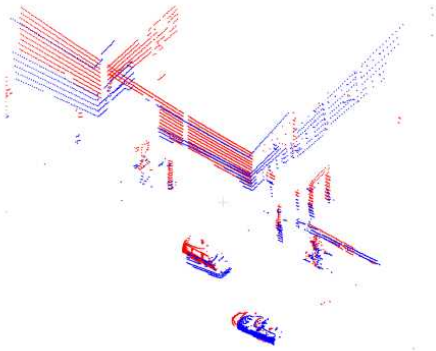
(a) Photo of environment (b) Distance between vehicles

Figure 5. Photo of environment and distance between vehicles (environment 2). The left photo is the bird's-eye view, and the yellow line in the left photo indicates the movement path of the vehicles.

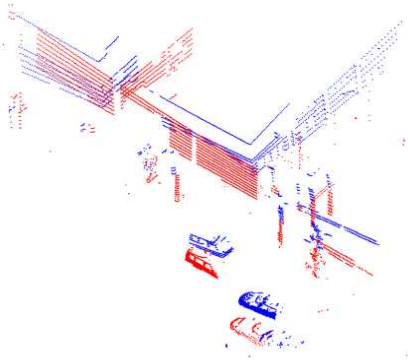
(a) Photo of environment



(b) Case 1 (using CSM)



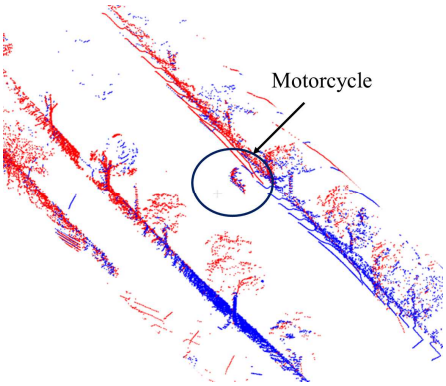
(c) Case 2 (using RTK GNSS)



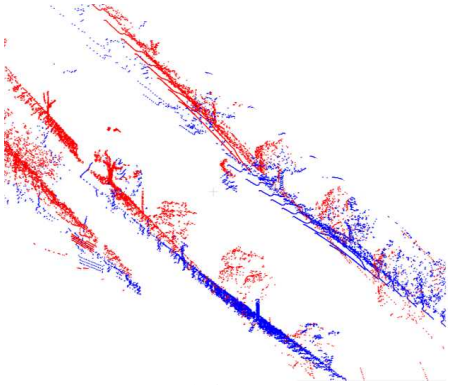
(d) Case 3 (using standard GNSS)

Figure 6. Mapping result in environment 1. The blue and red dots indicate the LiDAR measurements related to vehicles 1 and 2, respectively.

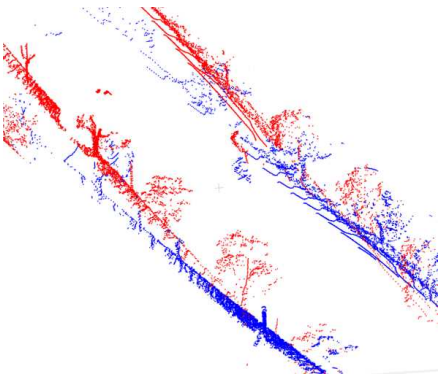
(a) Photo of environment



(b) Case 1 (using CSM)



(c) Case 2 (using RTK GNSS)



(d) Case 3 (using standard GNSS)

Figure 7. Mapping result in environment 2. The blue and red dots indicate the LiDAR measurements related to vehicles 1 and 2, respectively.

 ISBN: 978-1-61208-917-1

TABLE I. SUCCESS RATE OF CSM AND PERFORMANCE OF SUCCESS/FAILURE DECISION

		Environment 1	Environment 2
Success rate of CSM [%]		78.9	64.3
Performance of success/ failure decision	**Accuracy [%]**	95.5	87.4
	Precision [%]	97.8	86.1
	Recall [%]	96.6	99.0
	F-measure [%]	97.2	92.1

Table 1 also indicates that the performance of the CSM success/failure decision by the matching rate in environment 2 is considerably worse than that in environment 1. Since the road environment with street trees (environment 2) has many similar point feature histograms, incorrect matching frequently occurs. In consequence, the performance of the CSM success/failure decision is highly degraded in environment 2.

V. CONCLUSION AND FUTURE WORK

CMOT requires mapping LiDAR measurements captured by nearby vehicles onto a grid map represented on a common coordinate frame (e.g., world coordinate frame).

To accurately map these LiDAR measurements onto a grid map in GNSS-denied environments, this paper presented a relative pose estimation method that used CSM. Here, the relative pose between vehicles was estimated using the LiDAR measurements in the overlapping sensing areas of the vehicles. The relative pose estimation was performed using FPFH and RANSAC-based coarse registration and NDT scan matching-based fine registration. Experimental results using two LiDAR-equipped vehicles in two road environments showed that the proposed CSM has better applicability in urban environments with a high number of streets.

Since the spatial resolution of LiDAR is low in the vertical direction, the distance between vehicles where CSM can be achieved is short. Our current research effort aims to improve the CSM algorithm so that the relative pose can be estimated accurately even at long inter-vehicle distances. In addition, CSM will be implemented to CMOT and cooperative positioning.

REFERENCES

[1] E. Marti, J. Perez, M. A. Miguel, and F. Garcia, “A Review of Sensor Technologies for Perception in Automated Driving,” IEEE Intelligent Transportation Systems Magazine, pp. 94–108, 2019.

[2] F. P. Muller, “Survey on Ranging Sensors and Cooperative Techniques for Relative Positioning of Vehicles,” Sensors, vol. 17, 2017.

[3] E. Arnold, et al., “A Survey on 3D Object Detection Methods for Autonomous Driving Applications,” IEEE Trans. on Intelligent Transportation Systems, vol.20, no.10, pp. 3782–3795, 2019.

[4] Y. Tamura, R. Murabayashi, M. Hashimoto, and K. Takahashi, “Hierarchical Cooperative Tracking of Vehicles and People Using Laser Scanners Mounted on Multiple Mobile Robots,” Int. J. Advances in Intelligent Systems, vol. 10, pp. 90–101, 2017.

[5] Q. Chen, S. Tang, Q. Yangy, and S. Fuy, “Cooper: Cooperative Perception for Connected Autonomous Vehicles based on 3D Point Clouds,” Proc. of IEEE 39th Int. Conf. on Distributed Computing Systems (ICDCS), 2019.

[6] M. Shan, et al., “Demonstrations of Cooperative Perception: Safety and Robustness in Connected and Automated Vehicle Operations,” Sensors, 2021.

[7] A. Milstein, “Occupancy grid maps for localization and mapping. Motion Planning,” Xing-Jian Jing, ed., InTech, pp. 382–408, 2008

[8] 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.

[9] S. Kanaki, M. Hashimoto, Y. Yoden, and K. Takahashi, “Laser-based Cooperative Tracking of Vehicles and People by Multiple Mobile Robots in GNSS-denied Environments,” Proc. of IEEE Int. Conf. on Advanced Intelligent Mechatronics (AIM), pp. 1228–1233, 2017.

[10] G. Soatti, et al., “Implicit Cooperative Positioning in Vehicular Networks,” IEEE Trans. on Intelligent Transportation Systems, vol.19, issue 12, pp. 3964 –3980, 2018.

[11] X. Shen, et al., “A General Framework for Multi-Vehicle Cooperative Localization Using Pose Graph,” arXiv:1704.01252, 2017.

[12] P. J. Besl and N. D. McKay, “A Method of Registration of 3-D Shapes,” IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 14, no. 2, pp. 239–256, 1992.

[13] P. Biber and W. Strasser, “The Normal Distributions Transform: A New Approach to Laser Scan Matching,” Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS 2003), pp. 2743–2748, 2003.

[14] R. B. Rusu, N. Blodow, and M. Beetz, “Fast Point Feature Histograms (FPFH) for 3D Registration,” Proc. of IEEE/RSJ Int. Conf. on Robotics and Automation, pp. 3212–3217, 2009.

[15] A. Aldoma, et. al., “Tutorial: Point Cloud Library: Three-Dimensional Object Recognition and 6 DoF Pose Estimation,” IEEE Robotics & Automation Magazine, vol.19, issue 3, pp. 80–91, 2012.

[16] K. Inui, M. Morikawa, M. Hashimoto, and K. Takahashi, “Distortion Correction of Laser Scan Data from In-vehicle Laser Scanner based on Kalman Filter and NDT Scan Matching,” Proc. of the 14th Int. Conf. on Informatics in Control, Automation and Robotics (ICINCO), pp. 329–334, 2017.

[17] K. Tokorodani, M. Hashimoto, Y. Aihara, and K. Takahashi, “Point-Cloud Mapping Using Lidar Mounted on Two-Wheeled Vehicle Based on NDT Scan Matching,” Proc. of the 16th Int. Conf. on Informatics in Control, Automation and Robotics (ICINCO), pp.446–452, 2019.

[18] S. Kanaki, et al., “Cooperative Moving-Object Tracking with Multiple Mobile Sensor Nodes -Size and Posture Estimation of Moving Objects using In-vehicle Multilayer Laser Scanner-,” Proc. of 2016 IEEE Int. Conf. on Industrial Technology (ICIT 2016), pp. 59–64, 2016.

[19] S. Tanaka, C. Koshiro, M. Yamaji, M. Hashimoto, and K. Takahashi, “Point Cloud Mapping and Merging in GNSS-Denied and Dynamic Environments Using Only Onboard Scanning LiDAR,” Int. J. on Advances in Systems and Measurements, vol. 13 no. 3&4, pp. 275–288, 2020.

 ISBN: 978-1-61208-917-1