Toward Localization-Based Automated Driving in Highly Dynamic Environments: Comparison and Discussion of Observation Models

Naoki Akai[1], Luis Yoichi Morales[1], Takatsugu Hirayama[1], and Hiroshi Murase[2]

Abstract— To robustly localize the pose of an ego vehicle within a dynamic environment, it is important to model the sensor measurements precisely, including changes in the environment. This study describes the observation models developed for localization performed in highly dynamic environments, and presents the results of comparing these models. In this study, four observation models, including our previously proposed model, were compared by conducting a simulation. The models had different ways of coping with changes in the environment, and produced different results. Moreover, the comparison results revealed that each model had its own advantages and disadvantages. Finally, we demonstrated that our previously proposed model can achieve satisfactory performance in terms of computation complexity and estimation accuracy.

Keywords— Localization, Observation Model

I. INTRODUCTION

Vehicle self-localization, that is, a function to estimate an ego vehicle's relative pose on a given map, is fundamental for automated driving. An observation model, which represents sensor measurements probabilistically, is used to estimate the pose. Because it is difficult to model the dynamics of the environment, creating observation models that cope with environmental changes is also difficult. However, real environments change dynamically, owing to the movement of dynamic obstacles and landmarks, as shown in Fig. 1. Hence, it is not easy to perform robust localization in highly dynamic environments. This paper considers observation models to improve the robustness of localization against environmental changes.

A sensor, e.g., light detection and ranging (LiDAR), measures multiple data simultaneously, and the n-th sensor measurement data at time t, $\mathbf{y}_{nt}$ are generally formulated as follows:

$$\mathbf{y}_{nt} = h_{\mathrm{hit}}(\mathbf{x}_t, \mathbf{m}) + e_{nt}, \qquad (1)$$

where $\mathbf{x}_t$ is the pose of the ego vehicle, $\mathbf{m}$ is the given map, e_{nt} is the noise and/or error embedded in $\mathbf{y}_{nt}$, and $h_{\mathrm{hit}}(\cdot)$ is a function that returns the expected sensor measurements from the given pose. The error is basically modeled by a zero mean Gaussian distribution. However, it is impossible to cope with environmental dynamics when the error is modeled by a single Gaussian distribution, because such dynamics cannot be described by the single Gaussian distribution. Thus, the
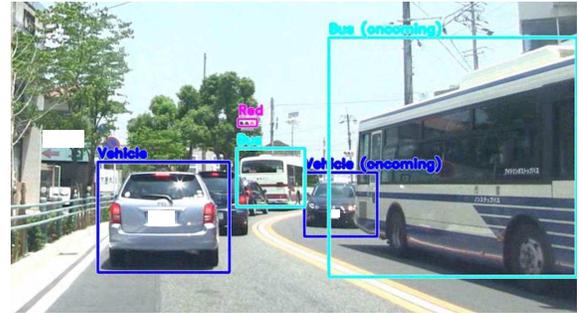
Fig. 1. Example of driving scene captured by camera mounted on ego vehicle. The exact modeling of sensor measurements includes changes in the environment, and this is significant for the robust localization of the ego vehicle's pose. This recognition result was reported by [1].

observation model is generally represented by a mixture of distributions [2], [3]. Particularly in the beam model [3], a probabilistic distribution with regard to observation against dynamic obstacles is included. Additionally, Thrun *et al.* has proposed a scan point rejection method, which might be reflected by dynamic obstacles. Thus, the robot pose can be estimated robustly, even if the robot is surrounded by many dynamic obstacles [2], [4]. However, this assumes that dynamic obstacles are equally disturbed in the environment and that the landmarks do not move. In most cases, these assumptions are not satisfied in real environments.

The sensor measurement error includes a component regarding the environment's shape. It might be considered that modeling the error in each area would improve the performance. Several approaches toward modelling the error by a mixture of simple distributions have been proposed, such as the normal distributions transform (NDT) [5] and the multiresolution Gaussian mixture model [6]. These approaches can flexibly model the error and improve the descriptive ability of Equation (1). However, these approaches do not correspond to high environmental dynamics because the environmental space is divided into several individual spaces, and the noise is modeled separately in each space.

Cao *et al.* proposed an approach toward error modelling by using a penalized mixture of exponential power (PMoEP) distributions [7]. The EP distribution is known as the generalized error distribution and includes all normal and Laplace distributions. Moreover, it includes all continuous uniform distributions on the bounded intervals of the real line, as limiting cases. Thus, the combination of EP distributions may be capable of modelling any error.

We have proposed an observation model that explicitly

[1]Naoki Akai, Luis Yoichi Morales, and Takatsugu Hirayama are with the Institute of Innovation for Future Society (MIRAI), Nagoya University, Nagoya 464-8601, Japan {akai, morales_yoichi}@coi.nagoya-u.ac.jp, takatsugu.hirayama@nagoya-u.jp

[2]Hiroshi Murase is with the Graduate School of Information Science, Nagoya University, Nagoya 464-8603, Japan murase@nagoya-u.jp

considers the environmental dynamics [8]. In this model, a single Gaussian distribution is used to model the error that exists in Equation (1). In this study , an observation model for describing the environmental dynamics was introduced. In short, the model calculates the probality, while considering whether sensor readings are obtained from known or unknown obstacles. Thus, it can cope with environmental dynamics.

This paper presents the comparison and evaluation results with regard to observation models constructed for localization in dynamic environments. Additionally, the details of these models are provided in this paper. The contribution of this study is that it provides concrete comparison results with regard to the observation models, and a discussion of these models, which were used to obtain the presented results. Moreover, the obtained results can be used to assist the selection of an observation model for the purpose of navigating dynamic environments.

The rest of this paper is organized as follows. Section II provides a summary of related studies. Section III describes the observation models used in this study. Section IV describes the implementation details. Section V presents the results of evaluating and comparing the models by conducting a simulation. Section VI presents the conclusions drawn from this study.

II. RELATED WORKS

To date, many types of localization approaches for autonomous cars have been proposed [10]–[15]. Various types of landmarks have been used for localization and automated driving has been achieved. Thus far, various researchers have assumed that the environment is static; however, this assumption is not suitable to real environments. Therefore, many researchers have tackled the localization problem in dynamic environments.

A straightforward approach consists of improving the observation model. In [2], an observation model that included both known and unknown obstacles was proposed. This model is known as a beam model [3] and is a combination of various distributions. A distribution represents the probability of measurements regarding unknown obstacles. To compute this probability, the ray tracing process, which calculates expected distances measurable from a current pose, is employed. One drawback of the beam model is the heavy computational workload required for ray tracing. The likelihood field model (LFM) has been proposed [3] to overcome this problem. This model is able to efficiently compute the probability. However, considering the influence of dynamic obstacles is not easy because the ray tracing process is omitted. Additionally, the beam and LFMs cannot cope with landmark removal.

Other observation models for coping with dynamic obstacles have been proposed [16], [17]. These models enable the robust calculation of probability against dynamic obstacles. However, it would be difficult to exactly estimate the ego vehicle pose in areas where landmarks do not exist in one side area. Yang *et al.* have proposed the feasibility

grids describing the dynamic environment facilitating the representation of information both for a stationary object and a moving object [18]. They introduced dual sensor models and modeled sensor measurements for moving obstacles by using feasibility grids. Our approach, which has been presented in [8], can cope both with dynamic obstacles and landmark removal, even if the ray tracing process is not employed. Additionally, our approach can simultaneously estimate the pose of the ego vehicle and dynamic obstacles without other information about the occupancy grid map, and without an increased computational cost. Cao *et al.* have proposed an error modeling approach toward low rank matrix factorization by using PMoEP distributions [7]. We introduce a method of using this approach with regard to a localization problem.

Another effective approach is to simultaneously estimate both the ego vehicle's pose and the environmental changes. Montemerlo *et al.* have proposed a simultaneous estimation approach toward pose and people tracking [19]. Approaches toward the simultaneous estimation of the pose and environment map in dynamic environments have also been proposed by several researchers [20]–[30]. Valencia *et al.* have proposed a localization approach that consisted of using a dual-timescale map, which enables the use of semi-static obstacles as landmarks [31]. These approaches allow autonomous cars to estimate their own pose by using a fresh map that accurately models the current environment. However, localization will fail when the map update process fails. Thus, the reliability of simultaneously estimating the pose and the environment map will be smaller than that of single localization, in terms of system integration. Our objective was to provide robust localization without the map update process, so as to maintain high reliability.

III. OBSERVATION MODELS

In this study, we compared four observation models. This section provides the description of each model.

A. Beam model and likelihood field model (LFM)

The beam model presented in [3] is expressed as follows:

$$p(\mathbf{y}_{nt}|\mathbf{x}_t, \mathbf{m}) = \begin{pmatrix} y_{\text{hit}} \\ y_{\text{short}} \\ y_{\text{max}} \\ y_{\text{rand}} \end{pmatrix}^T \cdot \begin{pmatrix} p_{\text{hit}}(\mathbf{y}_{nt}|\mathbf{x}_t, \mathbf{m}) \\ p_{\text{short}}(\mathbf{y}_{nt}|\mathbf{x}_t, \mathbf{m}) \\ p_{\text{max}}(\mathbf{y}_{nt}|\mathbf{x}_t, \mathbf{m}) \\ p_{\text{rand}}(\mathbf{y}_{nt}|\mathbf{x}_t, \mathbf{m}) \end{pmatrix},$$
(2)

where y_{hit}, y_{short}, y_{max}, and y_{rand} are arbitrary coefficients, whose sum must be equal to one. Additionally, p_{hit}, p_{short}, p_{max}, and p_{rand} are the probabilistic distributions of the measurement of known and unknown obstacles, measurement of maximum distance, and random measurement values, respectively. The details of each distribution can be found in [3].

To ascertain whether the measurement of the n-th sensor is obtained from known or unknown obstacles, the following

condition has been reported by [3]:

$$\frac{\int y_{\text{short}} p_{\text{short}}(\mathbf{y}_{nt}|\mathbf{x}_t, \mathbf{m})\overline{bel}(\mathbf{x}_t)d\mathbf{x}_t}{\int p(\mathbf{y}_{nt}|\mathbf{x}_t, \mathbf{m})\overline{bel}(\mathbf{x}_t)d\mathbf{x}_t} > \chi, \qquad (3)$$

where $\overline{bel}(\mathbf{x}_t)$ is the previous distribution regarding the pose at time t, and χ is the threshold. The measurement of the n-th sensor is rejected when the above condition is satisfied, and is not used to localize the pose. Practically, this equation is computed approximately by using a sampling-based method.

The beam model has the drawbacks of high computational cost and non-smooth probability distribution. To overcome these drawbacks, the LFM is proposed. The LFM is the mixture distribution of p_{hit}, p_{max}, and p_{rand}[1]. We evaluated the LFM with the scan points rejection method.

B. Normal distributions transform (NDT)

NDT [5] is an approximate representation method of point clouds with a set of normal distributions. In this method, the environment space is first divided into several spaces. Let $\mathbf{P}_i = [\mathbf{p}_{i0}, \mathbf{p}_{i1}, ..., \mathbf{p}_{iM}]^T$ be the points included in the i-th space; the mean $\mathbf{p}_i^m$ and covariance Σ_i are computed from these points. The mean and covariance set is called the ND map.

The probabilistic distribution regarding the measurement of known obstacles using the ND map is expressed as follows:

$$p_{\text{hit}}(\mathbf{y}_{nt}|\mathbf{x}_t, \mathbf{m}) = \\ \xi \exp \left\{ -\frac{1}{2} (\mathbf{p}_i^m - \mathbf{q}(\mathbf{y}_{nt}))^T \Sigma_i^{-1} (\mathbf{p}_i^m - \mathbf{q}(\mathbf{y}_{nt})) \right\}, \qquad (4)$$

where ξ is the normalization constant, $\mathbf{q}(\cdot)$ is the function that returns a point in the world coordinates corresponding to $\mathbf{y}_{nt}$; $\mathbf{p}_i^m$ and Σ_i are the mean point and covariance corresponding to the point returned by $\mathbf{q}(\cdot)$. We evaluate the NDT-based observation model by using the scan points rejection method. The p_{hit} of LFM is substituted in Equation (4) when the NDT-based observation model is used.

C. Conditional observation model (COM)

In our approach presented in [8], the sensor measurement class c_{nt} was introduced. Additionally, the pose of the ego vehicle and the class were estimated simultaneously. The probability distribution is expressed as follows:

$$p(\mathbf{y}_{nt}|\mathbf{x}_t, \mathbf{m}) = \int p(\mathbf{y}_{nt}|\mathbf{x}_t, c_{nt}, \mathbf{m})p(c_{nt})dc_{nt}, \quad (5)$$

where $p(\mathbf{y}_{nt}|\mathbf{x}_t, c_{nt}, \mathbf{m})$ is the observation model with a certain condition. Thus, it is called the conditional observation model (COM). In this study, the model considers two discrete classes, namely, $C = \{\text{known}, \text{unknown}\}$, and the probability distribution is expressed as follows:

$$p(\mathbf{y}_{nt}|\mathbf{x}_t, \mathbf{m}) = \\ p(\mathbf{y}_{nt}|\mathbf{x}_t, c_{nt} = \text{known}, \mathbf{m})p(c_{nt} = \text{known}) \\ + p(\mathbf{y}_{nt}|\mathbf{x}_t, c_{nt} = \text{unknown}, \mathbf{m})p(c_{nt} = \text{unknown}) \quad (6)$$

[1]In the beam and LFM, a slightly different p_{hit} is used. For details please refer to [3].

As expressed in the equation, we have two conditional observation models. Because the conditional observation model of $c_n = \text{known}$ is given, the LFM is used. The conditional observation model of $c_n = \text{unknown}$ is also given and expressed as follows :

$$p(\mathbf{y}_{nt}|\mathbf{x}_t, c_{nt} = \text{unknown}, \mathbf{m}) \\ = \xi \{\text{max}_{\text{known}} - p(\mathbf{y}_{nt}|\mathbf{x}_t, c_{nt} = \text{known}, \mathbf{m})\}, \qquad (7)$$

where $\text{max}_{\text{known}}$ is the maximum value of $p(\mathbf{y}_{nt}|\mathbf{x}_t, c_{nt} = \text{known}, \mathbf{m})$.

D. Penalized mixture of exponential power (PMoEP)

In the PMoEP model [7], the probabilistic distribution regarding the error is expressed as follows:

$$p(e_{nt}|\boldsymbol{\pi}_t, \boldsymbol{\eta}_t) \propto \sum_{k=1}^{K} \pi_{kt} f_{p_k}(e_{nt}; 0, \eta_{kt}), \qquad (8)$$

where $\boldsymbol{\pi}_t = [\pi_{1t}, \pi_{2t}, ..., \pi_{Kt}]^T$ is the proportion of mixing with $\pi_{kt} \geq 0$ and $\sum_{k=1}^{K} \pi_{kt} = 1$, K is the number of the mixture components, and $f_{p_k}(\cdot)$ is the k-th EP distribution. The k-th EP distribution $(p_k > 0)$ is expressed as follows:

$$f_{p_k}(e_{nt}; 0, \eta_{kt}) = \frac{p_k \eta_{kt}^{\frac{1}{p_k}}}{2\Gamma(\frac{1}{p_k})} \exp \left\{-\eta_{kt}|e_{nt}|^{p_k}\right\}, \qquad (9)$$

where $\boldsymbol{\eta}_t = [\eta_{1t}, \eta_{2t}, ..., \eta_{Kt}]^T$ is the precision parameter, p_k is the shape parameter, and $\Gamma(\cdot)$ is the Gamma function. By changing the shape parameter, the EP distribution can describe the leptokurtic $(0 < p < 2)$, mesokurtic $(p = 2)$, and platykurtic $(p > 2)$ distributions. In particular, the Laplace, Gaussian, and uniform distributions are obtained with $p = 1$, $p = 2$, and $p \to \infty$, respectively. Moreover, the error distribution is denoted as follows: $p(e_{nt}|\boldsymbol{\pi}_t, \boldsymbol{\eta}_t) = 2\sum_{k=1}^{K} \pi_{kt} f_{p_k}(e_{nt}; 0, \eta_{kt})$, because the error is defined only in the positive region.

The assumption that each error e_{nt} is equipped with an indicator variable $\mathbf{z}_{nt} = [z_{n1t}, z_{n2t}, ..., z_{nKt}]^T$, where $z_{nkt} \in \{0, 1\}$ and $\sum_{k=1}^{K} z_{nkt} = 1$, is introduced. $z_{nkt} = 1$ implies that the error e_{nt} is drawn from the k-th EP distribution, and that $\mathbf{z}_{nt}$ obeys a multinomial distribution. Thus, the following distributions can be obtained:

$$p(e_{nt}|\mathbf{z}_{nt}) = \prod_{k=1}^{K} f_{p_k}(e_{nt}; 0, \eta_{kt})^{z_{nkt}}, \qquad (10)$$

$$p(\mathbf{z}_{nt}) = \prod_{k=1}^{K} \pi_k^{z_{nkt}}. \qquad (11)$$

$\mathbf{E}_t = [e_{1t}, e_{2t}, ..., e_{Nt}]^T$, $\mathbf{Z}_t = [\mathbf{z}_{1t}, \mathbf{z}_{2t}, ..., \mathbf{z}_{Nt}]^T$, and $\boldsymbol{\Theta}_t = \{\boldsymbol{\pi}_t, \boldsymbol{\eta}_t\}$ are introduced, and the log-likelihood function regarding the errors can be expressed as follows:

$$\ln p(\mathbf{E}_t|\boldsymbol{\Theta}_t) = \ln \sum_{\mathbf{Z}_t} p(\mathbf{E}_t, \mathbf{Z}_t|\boldsymbol{\Theta}_t), \qquad (12)$$

where $p(\mathbf{E}_t, \mathbf{Z}_t | \boldsymbol{\Theta}_t)$ is the complete probability function and is expressed as follows:

$$p(\mathbf{E}_t, \mathbf{Z}_t | \boldsymbol{\Theta}_t) = \prod_{n=1}^{N} \prod_{k=1}^{K} \{\pi_{kt} f_{p_k}(e_{nt}; 0, \eta_{kt})\}^{z_{nkt}}. \quad (13)$$

Then, the complete log-likelihood function is expressed as follows:

$$\ln p(\mathbf{E}_t, \mathbf{Z}_t | \boldsymbol{\Theta}_t) =$$
$$\sum_{n=1}^{N} \sum_{k=1}^{K} z_{nkt} \{\ln \pi_{kt} + \ln f_{p_k}(e_{nt}; 0, \eta_{kt})\}. \quad (14)$$

In the PMoEP model, the penalized log-likelihood function is defined [9] and expressed as follows:

$$l(\boldsymbol{\Theta}_t) = \ln p(\mathbf{E}_t, \mathbf{Z}_t | \boldsymbol{\Theta}_t) - P(\boldsymbol{\pi}_t), \quad (15)$$

$$P(\boldsymbol{\pi}_t) = \lambda N D_f \sum_{k=1}^{K} \ln \frac{\epsilon + \pi_{kt}}{\epsilon}, \quad (16)$$

where ϵ is a very small positive number, λ is a tuning parameter ($\lambda > 0$), and D_f is the number of free parameters for the k-th component. In the PMoEP model used in this study, D_f was set to the value of two for π_{kt} and η_{kt}. The model was solved by an expectation-maximization (EM) algorithm, which implemented an iterative procedure. Additionally, we note that $\boldsymbol{\Theta}_t^{(i)} = \{\boldsymbol{\pi}_t^{(i)}, \boldsymbol{\eta}_t^{(i)}\}$ is the estimation at the i-th iteration.

In the E step, the responsibility γ_{nkt} is first updated on the basis of Bayes' theorem.

$$\gamma_{nkt}^{(i+1)} = \frac{\pi_{kt}^{(i)} f_{p_k}(e_{nt}; 0, \eta_{kt}^{(i)})}{\sum_{l=1}^{K} \pi_{lt}^{(i)} f_{p_l}(e_{nt}; 0, \eta_{lt}^{(i)})}. \quad (17)$$

Then, the Q function is formulated as follows:

$$Q(\boldsymbol{\Theta}_t, \boldsymbol{\Theta}_t^{(i)}) =$$
$$\sum_{n=1}^{N} \gamma_{nkt}^{(i+1)} \left\{ \ln f_{p_k}(e_{nt}; 0, \eta_{kt}^{(i)}) + \ln \pi_{kt}^{(i)} \right\} - P(\boldsymbol{\pi}_t^{(i)}). \quad (18)$$

In the M step, $\boldsymbol{\Theta}_t$ is updated by maximizing the Q function. The update equations are obtained by taking the first derivative of Q with respect to $\boldsymbol{\pi}_t$ and $\boldsymbol{\eta}_t$, and finding the zero points through the following equations:

$$\pi_{kt}^{(i+1)} = \max \left\{ 0, \frac{1}{1 - \lambda K D_f} \left(\frac{N_{kt}}{N} - \lambda D_f \right) \right\}, \quad (19)$$

$$\eta_{kt}^{(i+1)} = \frac{N_{kt}}{p_k \sum_{n=1}^{N} \gamma_{nkt}^{(i+1)} |e_{nt}|^{p_k}}, \quad (20)$$

$$N_{kt} = \sum_{n=1}^{N} \gamma_{nkt}^{(i+1)}. \quad (21)$$

With the EM algorithm, the probabilistic distribution regarding the $p(e_{nt} | \boldsymbol{\pi}_t, \boldsymbol{\eta}_t)$ error is obtained. Owing to the distribution model error expressed in Equation (1), the following relationship can be obtained:

$$p_{\text{hit}}(\mathbf{y}_{nt} | \mathbf{x}_t, \mathbf{m}) = p(e_{nt} | \boldsymbol{\pi}_t, \boldsymbol{\eta}_t). \quad (22)$$

When the PMoEP-based observation model is used, p_{hit} of LFM is substituted in Equation (22).

IV. IMPLEMENTATION

A. Probability calculation

Let $\mathbf{Y}_t = [\mathbf{y}_{1t}, \mathbf{y}_{2t}, ..., \mathbf{y}_{Nt}]^T$ be the sensor measurements at time t. Then, the probability distribution $p(\mathbf{Y}_t | \mathbf{x}_t, \mathbf{m})$ is calculated as follows:

$$p(\mathbf{Y}_t | \mathbf{x}_t, \mathbf{m}) = \prod_{n=1}^{N} p(\mathbf{y}_{nt} | \mathbf{x}_t, \mathbf{m}), \quad (23)$$

under the assumption that each sensor measurement is independent.

To calculate the probability by using the LFM and NDT, we first separate the sensor measurements by using Equation (3). Then, the probability is calculated by using only the accepted points, i.e., the points obtained from known obstacles. However, when the probability is calculated by using the methods described in subsection III-C and III-D, the rejection algorithm is not used, because these methods include a scheme for coping with environmental dynamics.

B. Sensor measurement simulation

In this study, we conducted a simulation to create two dimensional (2D) light detection and ranging (LiDAR) measurements. We used the UXM-30LAH-EWA scanning and laser rangefinder (HOKUYO AUTOMATIC CO., LTD.) [32] to simulate the 2D scan. This device has the following specifications: maximum range of 80 m, scanning angle of 190 deg , and scanning angle resolution of 0.125 deg.

C. Scan point rejection

To compute Equation (3), the prior distribution of the pose must first be defined. The pose consisted of a 2D point, i.e., x and y, and the heading direction θ. We approximated the prior distribution by using 200 particles with a Gaussian distribution, and its standard deviations for the x, y, and θ elements, were set to 0.1 m, 0.1 m, and 0.5 deg , respectively.

D. PMoEP

The precision parameter had to be provided in advance, and $\mathbf{p} = [0.5, 1, 1.5, 2, 5, 10, 20]^T$ was set. The maximum number of iterations in the EM algorithm was set to 20, and the turning parameter λ was set to the value of one.

V. COMPARISON AND DISCUSSION

A. Experimental conditions

To carry out the precise numerical evaluation of the observation models, we used a simulation environment. Figure 2 shows the experimental environment and its occupancy grid map. First, we drove a vehicle equipped with a LiDAR and applied the 2D grid SLAM reported by [33]. The resolution of the map was set to 0.1 m. Additionally, we built the ND map based on the SLAM result, and its resolution was set to 1 m.

We used two occupancy grid maps. The first map, which is called the simulation map, was used to simulate the 2D scan, while the second one, called the landmark map, was used to calculate the probability. To evaluate robustness
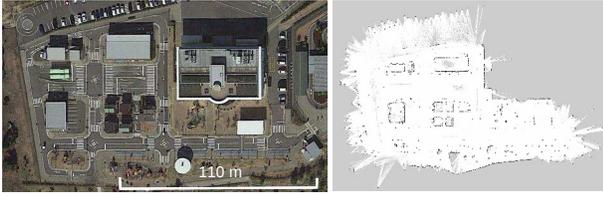
Fig. 2. Experimental environment (left) and its occupancy grid map (right).

against environmental changes, randomly moving dynamic obstacles were added to the simulation map. Additionally, we attempted to carry out verification in the case where the occupied grids were randomly removed from the simulation map. In this study, we evaluated four observation models; namely, LFM, COM, NDT-based models, and PMoEP-based models. In this section, these models are respectively noted as LFM, NDT, COM, and PMoEP.

B. Comparison

Figure 3 shows the comparison result, where the occupied grids were not removed from the simulation map. The top figure depicts the landmark map (black), and the laser scan points plotted from the ground truth pose (red). It should be noted that the simulation map was identical to the landmark map. The scan points that did not hit the landmarks depict the dynamic obstacles. The bottom figures show the probability distributions calculated by using each observation model around the ground truth (in $0, 0$). The probability values were normalized from 0 to 1. Δx and Δy are the differences in relation to the ground truth. In this comparison, the heading angle θ was not changed. As can be seen in the bottom figures, the probability distributions calculated with the COM and PMoEP created a peak around the ground truth.

Figure 4 shows the comparison result with approximately 95 % of the occupied grids having been removed from the simulation map at random. The top figure depicts the simulation map in green color. Even though the occupied grids were removed, the probability distributions calculated with the use of the COM and PMoEP also created a peak around the ground truth. Additionally, the probability distribution calculated by the PMoEP was more accurate than that of other probability distributions. These results revealed that the PMoEP could robustly and accurately estimate the pose of the ego vehicle in a dynamic environment. However, each model had a disadvantage. The observation models are discussed below.

C. Discussion

1) LFM: In the compared models, LFM is the simplest one. However, this model works in many cases. In particular, by using the scan point rejection method, a robust probability calculation against dynamic obstacles could be achieved. However, LFM was not robust against landmark removal because it did not have a describable model for the removal. Moreover, the scan point rejection method based on Equation (3) did not correspond to the removal because
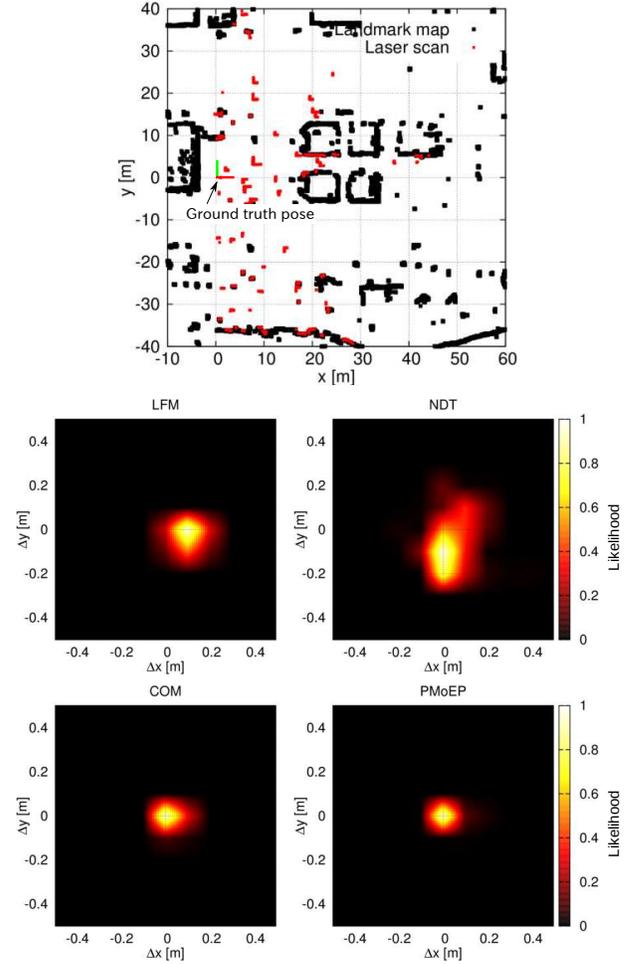
Fig. 3. Comparison result with occupied grids remaining in the simulation map.

it assumed that the landmarks were static. Therefore, LFM is not useful in environments where the landmarks move frequently. For the same reason, the NDT is also not useful in such environments.

Additionally, the performance of the scan point rejection method decreased when the prior distribution regarding the pose, i.e., the distribution of particles, was uncertain. To avoid a decrease in performance, the particles were clustered and the scan point rejection method was applied to each cluster. However, this leads to more complex problems, such as unsupervised clustering, as has been reported in [34].

2) NDT: As can be seen in Figs. 3 and 4, NDT exhibited the worst performance, in comparison with the rest of the models, owing to the low resolution of the ND map. Figure 5 shows the probability distributions calculated by the LFM and NDT with a different resolution ND map. We conducted this evaluation without landmark removal, because the scan point rejection method could not cope otherwise. However, the dynamic obstacles were set to the simulation map. When the resolution of the ND map was set to 0.1 m or 0.2 m, the observed performance was acceptable. However, the performance decreased with the decrease of resolution. The use of
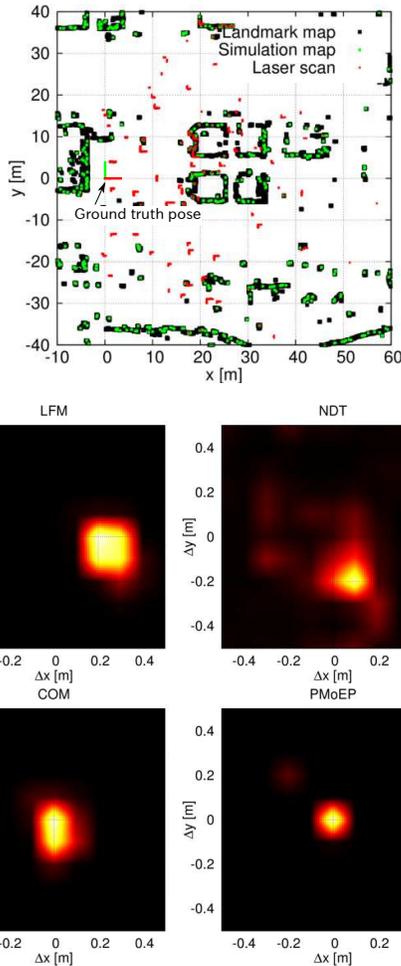
Fig. 4. Comparison result with occupied grids randomly removed from simulation map.

Fig. 6. Error modeling results in PMoEP. Top and bottom figures correspond to results shown in Figs. 3 and 4, respectively.

a very high-resolution map, e.g., 0.05 m, resulted in inferior performance, because it would lose the representability of suitable measurement errors.

The memory usage clearly increased when the high-resolution ND map was used. To obtain better performance than LFM, the resolution should be set to less than 0.5 m. However, this requires larger memory usage in comparison with that required by the occupancy grid map. The low resolution ND map enables the reduction of memory size, but results in decreased accuracy. Additionally, to update the ND map, a management process that is more complex than that of the occupancy grid map is required [35]. Finally, there exists a trade-off problem with regard to LFM and NDT.

3) COM: The COM provided good performance in the two comparison cases shown in Figs. 3 and 4. Even though the COM attempted to simultaneously estimate the pose of the ego vehicle and the class of the sensor measurements, its computational complexity was identical to that of the LFM because the class estimation and probability calculation could be carried out within the same computational loop. Additionally, the COM could cope both with the dynamic obstacles and landmark removal because a descriptive model

for the environmental changes was included.

A drawback of COM is that the model describing the changes in the environment is very imprecise. Ray tracing was used to calculate Equation (3). Thus, the existing probability for dynamic obstacles could be modeled by an exponential distribution. However, a uniform distribution was used to model the existing probability in the COM. Consequently, ray tracing could be omitted in the COM, and the time required for computation was reduced. Additionally, the COM corresponded to landmark removal. However, the COM could not produce a relatively high probability around the ground truth because it used the uniform distribution to represent the existing probability of dynamic obstacles. Therefore, the COM failed to localize the ego vehicle's pose when the initial guess regarding the pose was not accurate.

4) PMoEP: Figure 6 shows the error of the PMoEP results. The top and bottom figures correspond to the results shown in Figs. 3 and 4, respectively. The red and blue lines depict the modeling results corresponding to the maximum and minimum probabilities, while the green lines indicate the rest of the modeling results. As can be seen in the figures, various modeling results were obtained and it was demonstrated that the modeling result did not provide the maximum probability when the probability of zero error was the highest. In the case shown in Fig. 4, the modeling result provided the maximum probability when the probability of zero error was the lowest. These results revealed that suitable error modeling can improve the robustness of calculating the probability against changes in the environment.

Although PMoEP provided a fairly accurate estimation, it had the drawback of requiring heavy computation. Table I lists the number of iterations performed by the EM algorithm, and the computation time in ms. In this experiment, an Intel(R) Xeon(R) CPU E5–1650 v3@3.50 GHz was used,
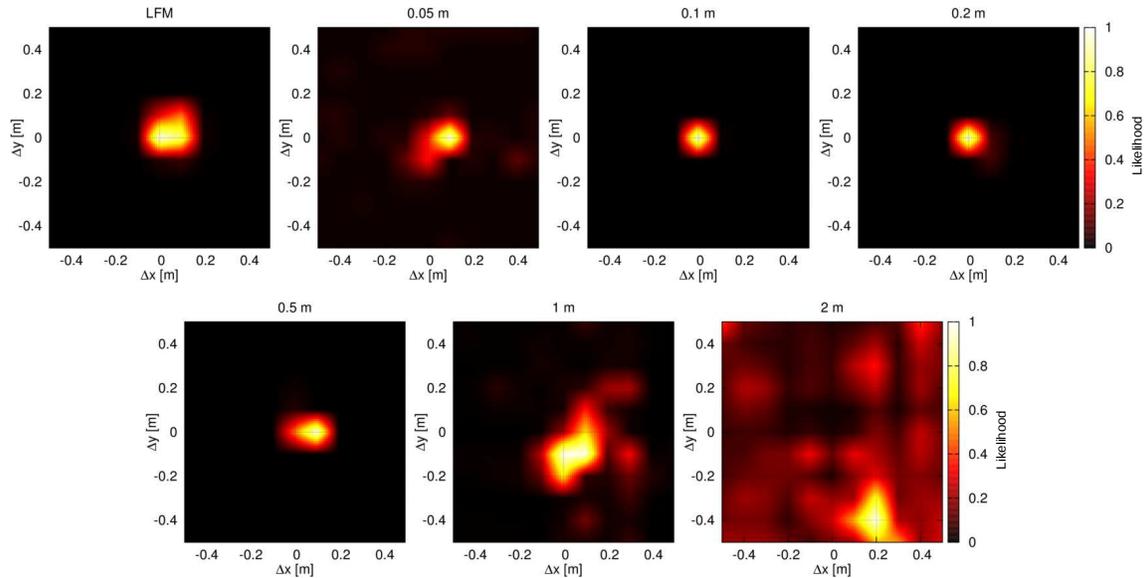
Fig. 5. Probability distributions calculated using LFM (top left) and NDT with different resolution ND map.

TABLE I

RESULTS REGARDING NUMBER OF ITERATIONS IN THE EM ALGORITHM AND COMPUTATION TIME (MS).

	Without landmark removal	With landmark removal
Min	1 (2.95 ms)	1 (2.92 ms)
Max	19 (44.46 ms)	18 (43.12 ms)
Ave	8.08 (19.68 ms)	7.56 (18.27 ms)
Std	4.65 (10.90 ms)	4.77 (11.05 ms)

Fig. 7. Number of iterations in the EM algorithm around ground truth.

TABLE II

SUMMARY OF COMPARISON RESULTS.

	Dynamic obstacles	Landmark removal	Memory usage	Computation time	Estimation accuracy in dynamic environments
LFM	◯	×	◯	◯	△
NDT	◯	×	△	◯	△
COM	◯	◯	◯	◯	◯
PMoEP	◯	◯	◯	×	◎

and the computation was carried out in a single thread. The results without and with landmark removal correspond to the results shown in Figs. 3 and 4. The mean value of the time to compute the probability with the PMoEP was approximately 20 ms. In our experience, the time required to compute the localization process must be less than 100 ms to achieve estimation in real-time. Therefore, using the PMoEP-based observation model for sampling-based localization is rather difficult. However, the rest of the compared models can be used for real-time estimation.

The distributions of the number of iterations around the ground truth are shown in Fig. 7. The number of iterations around the ground truth was not observed to be consistently smaller in comparison with that in the far areas. These results revealed that the computation time could not be reduced, even if the accuracy of the initial guess regarding the pose was increased. In other words, an acceleration method for the PMoEP is required for practical usage.

Finally, the comparison results are summarized in Table II. We concluded that the COM is suitable for localization in highly dynamic environments, in terms of computation complexity and estimation accuracy[2].

[2]The attached video shows the localization performance of the LFM and COM (https://www.youtube.com/watch?v=PIvBexeZcOw).

VI. CONCLUSION

This paper presented the results of comparing the observation models developed for localization in highly dynamic environments. In this study, four observation models were compared; namely, the LFM, NDT-based model, COM, and PMoEP-based model. The conclusions drawn from this study are as follows:

- The LFM may correspond to dynamic obstacles and work in many cases, but it cannot correspond to landmark removal and/or movement.
- The NDT cannot correspond to landmark removal, but can increase the estimation accuracy when a high resolution ND map is used. However, the memory usage increases as the resolution increases.
- The COM can correspond to both dynamic obstacles

and landmark removal. Additionally, its computational complexity is identical to that of the LFM. We concluded that the COM is suitable for localization in highly dynamic environments in terms of computational complexity and estimation accuracy.

- The PMoEP can accurately estimate a pose of an ego vehicle, even when the environment changes significantly. However, the PMoEP has a drawback in terms of computation time, and an acceleration method is required for use in practical situations.

We are also tackling a problem regarding the reliability estimation of a vehicle localization result to guarantee the localization performance [36]. The reliability estimation approach includes the error recognition of the localization result by using a machine learning approach. In future work, we will apply an exact observation model to error recognition.

REFERENCES

[1] D. Deguchi *et al*., "Intelligent traffic sign detector: Adaptive learning based on online gathering of training samples," in *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, pp. 72–77, 2011.

[2] D. Fox *et al*., "Markov localization for mobile robots in dynamic environments," *Journal of Artificial Intelligence Research (JAIR)*, vol. 11, pp. 391–427, 1999.

[3] S. Thrun *et al*., "Probabilistic robotics," *MIT Press*, 2005.

[4] W. Burgard *et al*., "Experiences with an interactive museum tour-guide robot," *Artificial Intelligence*, vol. 114, no. 1–2, 2000.

[5] P. Biber *et al*., "The normal distributions transform: A new approach to laser scan matching," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 3, 2003.

[6] R. W Wolcott *et al*., "Robust LIDAR localization using multiresolution Gaussian mixture maps for autonomous driving," *The International Journal of Robotics Research (IJRR)*, vol. 36, no. 3, pp. 292–319, 2017.

[7] X. Cao *et al*., "Low-rank matrix factorization under general mixture noise distributions," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 1493–1501, 2015.

[8] N. Akai *et al*., "Mobile robot localization considering class of sensor observations," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018 (accepted).

[9] T. Huang *et al*., "Model selection for Gaussian mixture models," *arXiv:1301.3558*, 2013.

[10] C. Brenner, "Global localization of vehicles using local pole patterns," in *Pattern Recognition. Springer* pp. 61–70, 2009.

[11] J. Levinson *et al*., "Robust vehicle localization in urban environments using probabilistic maps," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4372–4378, 2010.

[12] B. Qin *et al*., "Curb-intersection feature based Monte Carlo localization on urban roads," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2640–2646, 2012.

[13] A. Hata *et al*., "Road marker detection using LIDAR reflective intensity data and its application to vehicle localization," in *Proceedings of the IEEE International Conference on Intelligent Transportation Systems (ITSC)*, pp. 584–589, 2014.

[14] X. Qu *et al*., "Vehicle localization using mono-camera and geo-referenced traffic signs," in *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, 2015.

[15] N. Akai *et al*., "Autonomous driving based on accurate localization using multilayer LiDAR and dead reckoning," in *Proceedings of the IEEE International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1147–1152, 2017.

[16] S. Olufs *et al*., "An efficient area-based observation model for Monte-Carlo robot localization," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 13–20, 2009.

[17] E. Takeuchi *et al*., "A robust localization method based on free space observation model using 3D-map," in *Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 973–979, 2010.

[18] S. W. Yang *et al*., "Feasibility grids for localization and mapping in crowded urban scenes," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2011.

[19] M. Montemerlo *et al*., "Conditional particle filters for simultaneous mobile robot localization and people-tracking," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2002.

[20] D. Avots *et al*., "A probabilistic technique for simultaneous localization and door state estimation with mobile robots in dynamic environments," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2002.

[21] D. F. Wolf *et al*., "Mobile robot simultaneous localization and mapping in dynamic environments," *Autonomous Robots*, vol. 19, no. 1, pp. 53–65, 2005.

[22] L. Montesano *et al*., "Modeling the static and the dynamic parts of the environment to improve sensor-based navigation," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2005.

[23] C. Stachniss *et al*., "Mobile robot mapping and localization in non-static environments," in *Proceedings of the Association for the Advancement of Artificial Intelligence (AAAI)*, 2005

[24] A. Petrovskaya *et al*., "Probabilistic mobile manipulation in dynamic environments, with application to opening doors," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2007

[25] C. C. Wang *et al*., "Simultaneous localization, mapping and moving object tracking," *The International Journal of Robotics Research (IJRR)*, vol. 26, no. 9, pp. 889–916, 2007.

[26] S. Brechtel *et al*., "Recursive importance sampling for efficient grid-based occupancy filtering in dynamic environments," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2010.

[27] D. Meyer-Delius *et al*., "Temporary maps for robust localization in semi-static environments," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2010.

[28] D. Meyer-Delius *et al*., "Occupancy grid models for robot mapping in changing environments," in *Proceedings of the Association for the Advancement of Artificial Intelligence (AAAI)*, 2012.

[29] J. Saarinen *et al*, "Independent Markov chain occupancy grid maps for representation of dynamic environments," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2012.

[30] G. D. Tipaldi *et al*., "Lifelong localization in changing environments," *The International Journal of Robotics Research (IJRR)*, vol. 32, no. 14, pp. 1662–1678, 2013.

[31] R. Valencia *et al*., "Localization in highly dynamic environments using dual-timescale NDT-MCL," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2014.

[32] https://en.manu-systems.com/HOK-UXM-30LAH-EWA.shtml, accessed at July 16, 2018.

[33] G. Grisetti *et al*., "Improved techniques for grid mapping with rao-blackwellized particle filters," *IEEE Transactions on Robotics (TRO)*, V. 23, pp. 34–46, 2007.

[34] E. B. Fox *et al*., "The sticky hdp-hmm: Bayesian nonparametric hidden Markov models with persistent states," *Technical Report 2777, MIT Laboratory for Information and Decision Systems*, 2007.

[35] J. P. Saarinen *et al*., "3D normal distributions transform occupancy maps: An efficient representation for mapping in dynamic environments," *The International Journal of Robotics Research (IJRR)*, vol. 32, no. 14, pp. 1627–1644, 2013.

[36] N. Akai *et al*., "Reliability estimation of vehicle localization result," in *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, pp. 740–747, 2018.