

6DOF Iterative Closest Point Matching Considering A Priori with Maximum A Posteriori Estimation

Yoshitaka HARA[†], Shigeru BANDO[†], Takashi TSUBOUCHI[†],
Akira OSHIMA[‡], Itaru KITAHARA[†], Yoshinari KAMEDA[†]

[†] Graduate School of Systems and
Information Engineering,
University of Tsukuba, Japan

[‡] Doog Inc., Japan

Abstract— We present a new matching algorithm considering a priori (the prior probability) based on Bayes' theorem. Performance of point cloud registration between target and source clouds is effectively improved by introducing maximum a posteriori (MAP) estimation. The standard Iterative Closest Point (ICP) algorithm for the registration sometimes falls into misalignment due to measurement errors, narrow sensing field of view, or the movement of objects during measurement. Our approach resolves such problems by considering both the likelihood of the measurement and the prior probability of the initial guess for registration in the objective function. We have implemented a new 6DOF Iterative Closest Point matching using MAP estimation, and evaluated the method in real environments comparing with conventional registration methods. The experimental results have shown that our proposed method has wide convergence region and matches point clouds accurately preventing the misalignment problem.

I. INTRODUCTION

Autonomous navigation is an important research area, and the technology can be applied to cars, automated guided vehicles, construction machines, personal transporters, and mobile robots. Autonomous vehicles in DARPA Urban Challenge [1] and Google Self-Driving Car [2] are counting on accurate localization for navigation. In these works, 3D-LIDARs (laser scanners) are used primarily. On the other hand, various depth cameras that are inexpensive or even adapt to sunlight have appeared in the last few years.

Localization methods that match point clouds obtained from such sensors to map are often used. Conventional approaches for the registration of point clouds, however, sometimes fails to detect the ground truth and falls into misalignment due to measurement errors, narrow sensing field of view, or the movement of objects during measurement. These problems are exacerbated in 6 degrees of freedom (DOF) matching due to the significantly increased number of local minima, and hence matching is likely to fail. Bayesian approaches [3] are effective ways to approach the problem and many techniques have been proposed [4].

In this paper, we present a new matching algorithm introducing maximum a posteriori (MAP) estimation based on Bayes' theorem. A novel feature of this algorithm is to consider both the likelihood of the measurement and the prior probability (a priori) of the initial guess for registration in the objective function. This is a similar concept to the conventional methods using Bayesian estimation [4], but

solved with a different approach. We have implemented a new 6DOF Iterative Closest Point matching using MAP estimation, and evaluated localization errors of the method in comparison with conventional registration methods. We have named the proposed method MAP-ICP.

II. RELATED WORK

Point cloud matching has been enthusiastically researched in the computer vision field. [5] reviewed, in detail, coarse registration using features without an initial guess, and fine registration using the overall shapes and an initial guess. A representative method for fine registration that is the targeted in this paper is the Iterative Closest Point (ICP) algorithm [6], [7]. Besl et al. [6] have proposed point-to-point ICP, and Chen et al. [7] have proposed point-to-plane ICP. Segal et al. [8] have formulated a generalization of the point-to-point and point-to-plane, then introduced plane-to-plane ICP. Rusinkiewicz et al. [9] made comparative evaluations of ICP variants in terms of correspondence finding, weighting of correspondences, rejecting outliers, error metrics, and an optimization method. There are many researches using ICP variants for localization of mobile robots, such as [10]–[12]. Since the ICP algorithm is one of the least squares methods based only on the likelihood of the measurement, it sometimes falls into misalignment. ICP with M-estimation for robustness to outliers [13] and ICP with EM algorithm for the hypotheses of multiple correspondences [14] have been proposed, but these techniques are still a type of least squares method and have the problem described above.

To deal with the misalignment problem, our method takes into account both the likelihood of the measurement and the prior probability based on probability theory [3]. In robotics field, the Bayes Filter that is using Bayesian estimation is often used for localization [4]. The Extended Kalman Filter, that is an implementation of the Bayes Filter with the Gaussian distribution, however is not able to handle landmarks of arbitrary shapes and can only handle point landmarks and line landmarks [4], [15]. On the other hand, Histogram Filter and Particle Filter can handle point clouds of arbitrary shapes as landmarks with grid representation [4]. Olson [16] has done real time 3DOF matching for 2D shapes by speeding up the localization method using correlation and Histogram Filter [17]. However, applying this method to

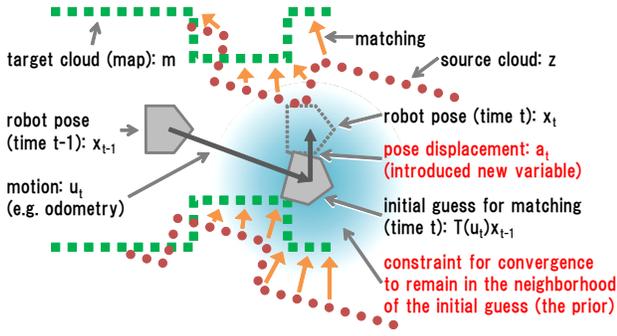


Fig. 1. A concept of the point cloud matching using maximum a posteriori estimation and the definitions of variables

3D shapes and 6DOF matching still requires an intractable computational cost. Kummerle et al. [18] and Suzuki et al. [19] applied Monte Carlo Localization [20] using a Particle Filter to 3D shapes and 6DOF matching. Our objective is also 6DOF matching with 3D shapes based on probability theory, but we perform the matching using MAP estimation instead of conventional Bayesian estimation. We have evaluated the performance of our method in real environments comparing with conventional registration methods.

III. MATCHING WITH MAP ESTIMATION

In this Section, we describe our algorithm for matching using maximum a posteriori estimation.

A. Formulation

The essence of the proposed method is considering both the likelihood of the measurement and the prior probability of the initial guess for registration in the objective function. For example, we use odometry as the prior probability of the motion. In the standard ICP, the odometry is conventionally used merely as the initial guess of the iterative calculation, and the algorithm is to look for a matching pose in the evaluation field represented only by the likelihood of the measurement. The proposed method finds a matching pose in the evaluation field represented both by the likelihood of the measurement and the prior probability.

Fig. 1 shows the concept of the matching using MAP estimation and definitions of respective variables. The proposed method matches point clouds under the constraint that they must remain in the neighborhood of the initial guess as the prior probability. In addition, it estimates the pose displacement in the coordinate frame where the origin is at the initial guess, in order to deal with the prior probability as a Gaussian distribution with zero mean. That is to say, it does not estimate robot pose x_t directly, it estimates the pose displacement a_t from the initial guess $T(u_t)x_{t-1}$. The initial guess $T(u_t)x_{t-1}$ is transformed robot pose x_{t-1} at time $t-1$ by motion u_t . Eq. (1) shows the relationship of respective variables.

$$\hat{x}_t = T(\hat{a}_t)T(u_t)x_{t-1} \quad (1)$$

Here, \hat{x}_t is the estimated robot pose x_t at time t , \hat{a}_t is the estimated pose displacement a_t at time t , $T(\hat{a}_t)$ is the

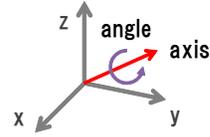


Fig. 2. Axis-angle representation for rotation in 3D Euclidean space

transformation matrix of the estimated pose displacement \hat{a}_t , and $T(u_t)$ is the transformation matrix of motion u_t .

For MAP estimation, we derive an energy function based on Bayes' theorem and perform the matching process by minimizing the function. The derived energy function $E(a_t)$ is shown in Eq. (2). The derivation of this equation will appear in Section III-D.

$$E(a_t) = \frac{1}{K} \sum_{k=1}^K \left\| T(a_t)z_{t,k} - m_{c_{t,k}} \right\|^2 + a_t^T \psi a_t \quad (2)$$

$z_{t,k}$ is the k -th point of the measured cloud (source cloud) at time t , K is the point number of the source cloud $z_{t,k}$, $m_{c_{t,k}}$ is the point of the map (target cloud) corresponding to source cloud $z_{t,k}$, ψ is a weighting matrix called a regularization parameter matrix. Compared with the error function of the standard ICP algorithm, the second term $+a_t^T \psi a_t$ is newly introduced. With the second term, the convergence will be constrained to remain in the neighborhood of the initial guess as the prior probability. The second term is equivalent to the regularization term (penalty term) according to [3]. The regularization parameter matrix ψ , that is described in detail in Section III-D, is defined by Eq. (3) using the covariance matrix of the prior probability Σ_a , the variance of the likelihood of the measurement σ_z^2 , and the point number of the source cloud K .

$$\psi = \frac{\sigma_z^2}{K} \Sigma_a^{-1} \quad (3)$$

Here, we multiply the variance of the likelihood of the measurement σ_z^2 by $\frac{1}{K}$ so that the ratio of the likelihood of the measurement (the first term) and the regularization term (the second term) in Eq. (2) does not depend on the point number of the source cloud K . Thus, the weighting of the likelihood of the measurement is invariant to the point number of the source cloud K .

The regularization term in Eq. (2) can be interpreted as the Mahalanobis distance of the pose displacement a_t with the regularization parameter matrix ψ . Here, we should pay attention to treatment of the rotation in 3D Euclidean space. For the regularization term, the distance (i.e. the norm) in not only translation but also rotation must be defined. There is the significant issue of dealing with the norm of 3DOF rotation of the pose displacement a_t . The Euler angles are not suitable to obtain the norm because of the ambiguity and the nonlinearity of the rotation representation. Additionally, the determinant of a rotation matrix and the norm of a rotation quaternion are always 1. Hence, we have selected the rotation angle in the axis-angle representation [21] as the rotation element of the norm. Fig. 2 shows a visualization of

a rotation represented by the axis-angle representation. The rotation angle value is equivalent to the norm of the rotation vector [21]. Finally the regularization parameter matrix ψ becomes a 4×4 matrix with 3 translation components and 1 rotation component.

The pose displacement \mathbf{a}_t that minimizes the energy function $E(\mathbf{a}_t)$ becomes the guess of displacement $\hat{\mathbf{a}}_t$ as shown in Eq. (4). We compute the guess of robot pose $\hat{\mathbf{x}}_t$ by substituting the estimated pose displacement $\hat{\mathbf{a}}_t$ into Eq. (1).

$$\hat{\mathbf{a}}_t = \underset{\mathbf{a}_t}{\operatorname{argmin}} E(\mathbf{a}_t) \quad (4)$$

B. MAP-ICP Algorithm

In order to minimize as in Eq. (4), it is necessary to know the correspondences from source cloud $\mathbf{z}_{t,k}$ to target cloud $\mathbf{m}_{c_{t,k}}$. In this paper we aim for fine registration using the overall shapes, our MAP-ICP algorithm determines the corresponding point by the nearest neighbor search and performs minimization of Eq. (4) in iterative calculation of the three steps that is a variant of the standard ICP algorithm. As described before, MAP-ICP computes the steps in the coordinate frame where the origin is at the initial guess of the iterative calculation, in order to deal with the prior probability as a Gaussian distribution with zero mean.

1) *Correspondence by the Nearest Neighbor Search:* In the first step, the algorithm searches for a target point $\mathbf{m}_{\hat{c}_{t,k}^{(i)}}$ on the map where the correspondence distance is minimum, relative to the respective source points $\hat{\mathbf{z}}_{t,k}^{(i-1)}$ in the previous iteration. Here (i) is the number of iterations $(1, 2, \dots, I)$, $\hat{c}_{t,k}^{(i)}$ is an estimated index of the target point \mathbf{m} corresponding to the source point $\hat{\mathbf{z}}_{t,k}^{(i-1)}$. $\hat{\mathbf{z}}_{t,k}^{(i-1)}$ is calculated in the previous iteration, where $\hat{\mathbf{z}}_{t,k}^{(0)} = \mathbf{z}_{t,k}$ (source point at the initial guess).

2) *Estimation of Differential Transformation:* Under the correspondences $\hat{c}_{t,k}^{(i)}$, the algorithm calculates the differential transformation $\hat{\mathbf{b}}_t^{(i)}$ from the previous pose displacement $\hat{\mathbf{a}}_t^{(i-1)}$ by minimizing as shown in Eq. (5) and (6).

$$\hat{\mathbf{b}}_t^{(i)} = \underset{\mathbf{b}_t^{(i)}}{\operatorname{argmin}} \left\{ \frac{1}{K} \sum_{k=1}^K \left\| \mathbf{T} \left(\mathbf{b}_t^{(i)} \right) \hat{\mathbf{z}}_{t,k}^{(i-1)} - \mathbf{m}_{\hat{c}_{t,k}^{(i)}} \right\|^2 + \mathbf{a}_t^{(i)T} \psi \mathbf{a}_t^{(i)} \right\} \quad (5)$$

$$\mathbf{a}_t^{(i)} = \mathbf{T} \left(\hat{\mathbf{b}}_t^{(i)} \right) \hat{\mathbf{a}}_t^{(i-1)} \quad (6)$$

We use Levenberg-Marquardt algorithm for nonlinear optimization shown in Eq. (5).

3) *Applying the Differential Transformation and Convergence Determination:* The algorithm transforms the pose displacement $\hat{\mathbf{a}}_t^{(i-1)}$ by the estimated differential transformation $\hat{\mathbf{b}}_t^{(i)}$ for the pose displacement $\hat{\mathbf{a}}_t^{(i)}$ in the i -th iteration as $\hat{\mathbf{a}}_t^{(i)} = \mathbf{T} \left(\hat{\mathbf{b}}_t^{(i)} \right) \hat{\mathbf{a}}_t^{(i-1)}$, where $\hat{\mathbf{a}}_t^{(0)} = \mathbf{0}$. Then, it transforms the respective source points $\hat{\mathbf{z}}_{t,k}^{(i-1)}$ by the estimated differential transformation $\hat{\mathbf{b}}_t^{(i)}$ as $\hat{\mathbf{z}}_{t,k}^{(i)} = \mathbf{T} \left(\hat{\mathbf{b}}_t^{(i)} \right) \hat{\mathbf{z}}_{t,k}^{(i-1)}$. If the differential transformation $\hat{\mathbf{b}}_t^{(i)}$ has converged to a sufficiently small

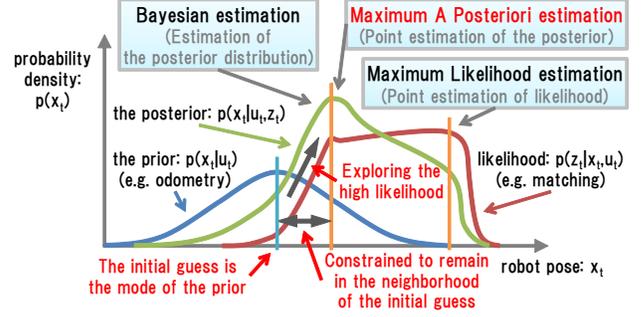


Fig. 3. A schematic graph of probability distribution in Bayes' theorem for localization

value, or if the number of iterations (i) reaches the threshold, the iterative calculation of MAP-ICP will terminate, otherwise it will return to the first step.

As a result of the above iterative calculation, the energy function $E(\mathbf{a}_t)$ in Eq. (2) is minimized and the source cloud $\mathbf{z}_{t,k}$ is matched to the target cloud $\mathbf{m}_{c_{t,k}}$. The algorithm also computes the guess of robot pose $\hat{\mathbf{x}}_t$ by substituting the estimated pose displacement $\hat{\mathbf{a}}_t^{(i)}$ into Eq. (1).

C. Comparison with Localization Methods Using Bayesian Estimation

In probability theory according to [3], the meaning of our proposed method is as follows. Applying Bayes' theorem to fundamental localization problem, we get Eq. (7).

$$p(\mathbf{x}_t | \mathbf{u}_t, \mathbf{z}_t) = \eta p(\mathbf{z}_t | \mathbf{x}_t, \mathbf{u}_t) p(\mathbf{x}_t | \mathbf{u}_t) \quad (7)$$

Here, $p(\mathbf{x}_t | \mathbf{u}_t, \mathbf{z}_t)$ is the posterior probability of the robot pose, $p(\mathbf{z}_t | \mathbf{x}_t, \mathbf{u}_t)$ is the likelihood of the measurement, $p(\mathbf{x}_t | \mathbf{u}_t)$ is the prior probability such as the odometry, and η is a normalization factor. Fig. 3 shows a schematic graph of probability density in the case of 1DOF localization for Eq. (7). The maximum likelihood estimation (equivalent to the least squares method) such as in the standard ICP obtains a guess by maximizing the likelihood. The estimation is based only on the likelihood of the measurement, it sometimes falls into misalignment. In contrast, MAP estimation and Bayesian estimation estimates the posterior probability. Taking account of both the likelihood of the measurement and the prior probability, it means the estimation is less likely to fall into misalignment. Our proposed method explores the guess of robot pose that is not too far from the mode of the prior probability and has high likelihood by using the MAP estimation.

In the Bayes Filter that is a type of Bayesian estimation for localization, we apply Bayes' theorem as in Eq. (8). Then, we get Eq. (9) by assuming the time series of robot poses as a Markov process and using the law of total probability. Eq. (9) shows the Bayes Filter for localization.

$$\begin{aligned} & p(\mathbf{x}_t | \mathbf{u}_{1:t}, \mathbf{z}_{1:t}, \mathbf{m}) \\ &= \eta p(\mathbf{z}_t | \mathbf{x}_t, \mathbf{u}_{1:t}, \mathbf{z}_{1:t-1}, \mathbf{m}) p(\mathbf{x}_t | \mathbf{u}_{1:t}, \mathbf{z}_{1:t-1}, \mathbf{m}) \quad (8) \\ &= \eta p(\mathbf{z}_t | \mathbf{x}_t, \mathbf{m}) \\ & \cdot \int p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{m}) p(\mathbf{x}_{t-1} | \mathbf{u}_{1:t-1}, \mathbf{z}_{1:t-1}, \mathbf{m}) d\mathbf{x}_{t-1} \quad (9) \end{aligned}$$

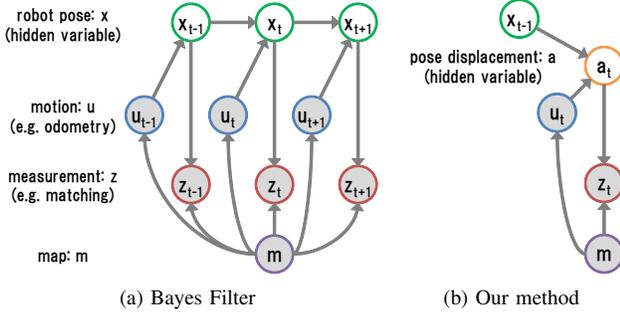


Fig. 4. Graphical models for localization

Localization is represented by the graphical models [3] as shown in Fig. 4. Fig. 4(a) is Bayes Filter, and Fig. 4(b) is our proposed method using MAP estimation. While Bayes Filter applies Bayes' theorem by assuming the time series of robot poses as a Markov process, the proposed method applies Bayes' theorem by dealing with the issue of localization at a certain point in time rather than the time series of robot poses. In addition, our method introduces the pose displacement as a new hidden variable in order to deal with the prior probability as a Gaussian distribution with zero mean. As shown in Fig. 3, Bayes Filter computes the posterior probability of the robot pose as a distribution, in comparison to our method that is one of the point estimation and computes only the mode (i.e. the best guess) of the posterior probability. If we require the distribution of the posterior probability, we can obtain it by the Laplace approximation [3].

D. Derivation of the Energy Function

In this Section, we describe the derivation of the energy function $E(\mathbf{a}_t)$ shown in Eq. (2). As mentioned before, the proposed method formulates the issue of localization at a certain point in time by introducing the pose displacement. Applying Bayes' theorem with the graphical model as shown in Fig. 4(b), we get Eq. (10).

$$p(\mathbf{a}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{z}_t, \mathbf{m}) = \eta p(\mathbf{z}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{a}_t, \mathbf{m}) p(\mathbf{a}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{m}) \quad (10)$$

For MAP estimation, we estimate the pose displacement \mathbf{a}_t as in Eq. (11) by maximizing the posterior probability in Eq. (10).

$$\hat{\mathbf{a}}_t = \underset{\mathbf{a}_t}{\operatorname{argmax}} p(\mathbf{a}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{z}_t, \mathbf{m}) \quad (11)$$

Assuming the likelihood distribution as the Gaussian distribution of the distances between corresponding points, the likelihood of the measurement becomes Eq. (12).

$$p(\mathbf{z}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{a}_t, \mathbf{m}) = \prod_{k=1}^K \mathcal{N}(\mathbf{z}'_{t,k} | S(\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{a}_t, \mathbf{m}, \mathbf{z}_{t,k}), \Sigma_z) \quad (12)$$

Here, $S(\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{a}_t, \mathbf{m}, \mathbf{z}_{t,k})$ is a function of the nearest neighbor search, and $\mathbf{z}'_{t,k}$ is the transformed source cloud $\mathbf{z}_{t,k}$ by the pose displacement \mathbf{a}_t as $\mathbf{z}'_{t,k} = \mathbf{T}(\mathbf{a}_t)\mathbf{z}_{t,k}$.

In addition, assuming that the covariance matrix of the likelihood distribution Σ_z is an isotropic covariance matrix [3] (a scalar matrix), we get $\Sigma_z = \sigma_z^2 \mathbf{I}$. Then, Eq. (12) becomes Eq. (13), where M is the dimension of the source cloud $\mathbf{z}_{t,k}$. In the case of 3D shapes, M is 3.

$$p(\mathbf{z}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{a}_t, \mathbf{m}) = \prod_{k=1}^K \frac{1}{(\sqrt{2\pi\sigma_z^2})^M} \cdot \exp\left\{-\frac{1}{2\sigma_z^2} \|\mathbf{z}'_{t,k} - S(\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{a}_t, \mathbf{m}, \mathbf{z}_{t,k})\|^2\right\} \quad (13)$$

Then, assuming the prior probability of the pose displacement as a Gaussian distribution with zero mean, we get

$$p(\mathbf{a}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{m}) = \mathcal{N}(\mathbf{a}_t | \mathbf{0}, \Sigma_a) \quad (14)$$

Eq. (14) is expanded to Eq. (15), where N is the dimension of the pose displacement \mathbf{a}_t . In the case of the displacement in 3D Euclidean space, N is 6DOF.

$$p(\mathbf{a}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{m}) = \frac{1}{(\sqrt{2\pi})^N} \frac{1}{\sqrt{|\Sigma_a|}} \exp\left(-\frac{1}{2} \mathbf{a}_t^T \Sigma_a^{-1} \mathbf{a}_t\right) \quad (15)$$

Since maximization of the posterior probability as shown in Eq. (11) is equivalent to minimizing the negative natural logarithm of the posterior probability, as

$$\hat{\mathbf{a}}_t = \underset{\mathbf{a}_t}{\operatorname{argmin}} -\ln p(\mathbf{a}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{z}_t, \mathbf{m}) \quad (16)$$

According to Eq. (16), we substitute Eq. (13) and Eq. (15) for negative natural logarithm of Eq. (10), then

$$\begin{aligned} & -\ln p(\mathbf{a}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{z}_t, \mathbf{m}) \\ &= -\ln p(\mathbf{z}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{a}_t, \mathbf{m}) - \ln p(\mathbf{a}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{m}) - \text{const.} \\ &= \frac{MK}{2} \ln(2\pi\sigma_z^2) + \frac{1}{2\sigma_z^2} \sum_{k=1}^K \|\mathbf{z}'_{t,k} - S(\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{a}_t, \mathbf{m}, \mathbf{z}_{t,k})\|^2 \\ & \quad + \frac{N}{2} \ln(2\pi) + \frac{1}{2} \ln|\Sigma_a| + \frac{1}{2} \mathbf{a}_t^T \Sigma_a^{-1} \mathbf{a}_t - \text{const.} \end{aligned} \quad (17)$$

We define the energy function $E(\mathbf{a}_t)$ shown in Eq. (18) by eliminating constant terms from Eq. (17) and multiplying it by $\frac{2\sigma_z^2}{K}$. As described in Section III-A, the regularization parameter matrix Ψ is defined by Eq. (3), which incorporated the covariance matrix of the prior probability Σ_a , the variance of the likelihood of the measurement σ_z^2 , and the point number of the source cloud K . Furthermore we expand the transformation of the transformed source cloud $\mathbf{z}'_{t,k}$ and the function of the nearest neighbor search $S(\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{a}_t, \mathbf{m}, \mathbf{z}_{t,k})$, finally we get Eq. (19).

$$E(\mathbf{a}_t) = \frac{1}{K} \sum_{k=1}^K \|\mathbf{z}'_{t,k} - S(\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{a}_t, \mathbf{m}, \mathbf{z}_{t,k})\|^2 + \mathbf{a}_t^T \Psi \mathbf{a}_t \quad (18)$$

$$= \frac{1}{K} \sum_{k=1}^K \|\mathbf{T}(\mathbf{a}_t)\mathbf{z}_{t,k} - \mathbf{m}_{c_{t,k}}\|^2 + \mathbf{a}_t^T \Psi \mathbf{a}_t \quad (19)$$

By the above calculation, we have derived the energy function $E(\mathbf{a}_t)$ of the proposed method shown in Eq. (2) based on Bayes' theorem.

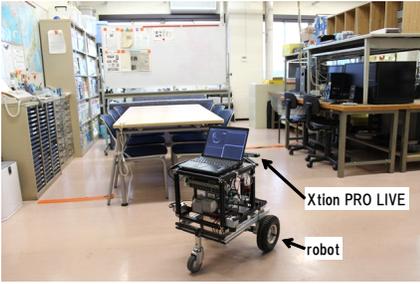


Fig. 5. An experimental environment to investigate the influence of complex shapes



Fig. 6. An experimental environment to investigate the influence of planar shapes

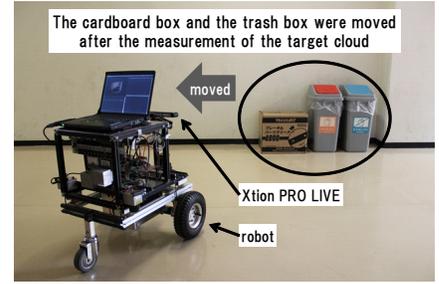


Fig. 7. An experimental environment to investigate the influence of moved objects

IV. EXPERIMENTS

In order to evaluate the proposed method MAP-ICP in comparison with standard ICP, we have experimented with a robot equipped with a depth camera ASUS Xtion PRO LIVE. Xtion has large measurement errors according to measuring distance because it uses the triangulation method. In addition, the sensing field of view of Xtion is narrow only 58 deg horizontally and 45 deg vertically, hence the registration is prone to fall into misalignment. We performed MAP-ICP and standard ICP under many experimental conditions of initial guesses, and evaluated localization errors of MAP-ICP in comparison with conventional standard ICP. For experimental conditions of initial guesses, we have added several translation and orientation errors independently to the ground truth. In this paper, the added translation errors are in the direction of the y axis (the lateral direction of the robot), and orientation errors are in the yaw rotation. In order to evaluate localization errors, we have used the distance from the ground truth for translation errors and the rotation angle from the ground truth in the axis-angle representation [21] for orientation errors.

It is known that the performance of the ICP algorithm will change depending on the definition of the error metric [9] (the norms in the first term of Eq. (2) and (5)). Therefore, we have experimented with both point-to-point [6] and point-to-plane [7] versions of matching. For point-to-plane matching, we compute the normals using the eigen decomposition of the covariance matrix of the closest points within 0.2 m of respective target points. The eigen vector corresponding to the smallest eigen value will approximate the surface normal. For both the point-to-point and the point-to-plane matching, we compare the four registration methods: standard ICP, MAP-ICP, standard ICP with RANSAC [22], and MAP-ICP with RANSAC. We use the implementation of Point Cloud Library [23] for standard ICP and standard ICP with RANSAC. In all of these algorithms, the maximum correspondence distance (threshold) was set to 1.0 m. The RANSAC outlier rejection threshold was set to 0.2 m. For MAP-ICP in this experiment, we assume the regularization parameter matrix ψ to be a diagonal matrix in that x and y elements were set to $\exp(-100)$, z element was set to $\exp(-5)$, and the axis-angle element was set to $\exp(-3)$. The x and y elements are very weak constraints. Here, the unit for translation components of the regularization parameter

matrix ψ is meter, and the unit for rotation component is radian.

We chose three experimental environments. The first is a room environment to investigate the influence of complex shapes where various objects such as desks and chairs are located as shown in Fig. 5. The second is a hallway environment that consists of large planes to investigate the influence of planar shapes as shown in Fig. 6. The third is a dynamic environment to investigate the influence of moved objects where some objects has been moved by us between the measurement of the target and source clouds as shown in Fig. 7. After the measurement of the target cloud, we have moved the cardboard box and the trash box to the left about 0.5 m. Then, the source cloud has been measured.

Fig. 8, 11, and 14 show examples of the matching results for standard ICP and MAP-ICP (without RANSAC) in each environment. In these examples, the initial guesses were the ground truths (the added errors are zeros) and point-to-plane matching was performed. Note that black means the target cloud, and gray means the source cloud. In the room environment as shown in Fig. 8, both standard ICP and MAP-ICP matched the source cloud accurately to the target cloud. Standard ICP, however, fell into misalignment in the hallway environment as shown in Fig. 11, while MAP-ICP was able to match point clouds accurately and keep the horizontal and the vertical tilt of the floor and the walls. In addition, standard ICP matched point clouds inaccurately under the influence of the moved objects in the dynamic environment as shown in Fig. 14, while MAP-ICP was still able to match point clouds accurately with constraint of the prior probability.

In each plot of Fig. 9, 10, 12, 13, 15, and 16, the vertical axis represents localization error where a lower value is a better result. We consider the matching has converged accurately in the following cases: the translation error is not more than 0.2 m, and the orientation error is not more than 5 deg. The horizontal axis represents the initial displacement from the ground truth (i.e. added error of the initial guess). The convergence region refers to the range of initial displacement to converge accurately.

Fig. 9 and 10 show the plots of localization error as a function of the initial displacement in the room environment. Fig. 9 shows point-to-point matching, while Fig. 10 shows point-to-plane. In both the point-to-point and the point-to-plane cases, MAP-ICP with RANSAC had the widest

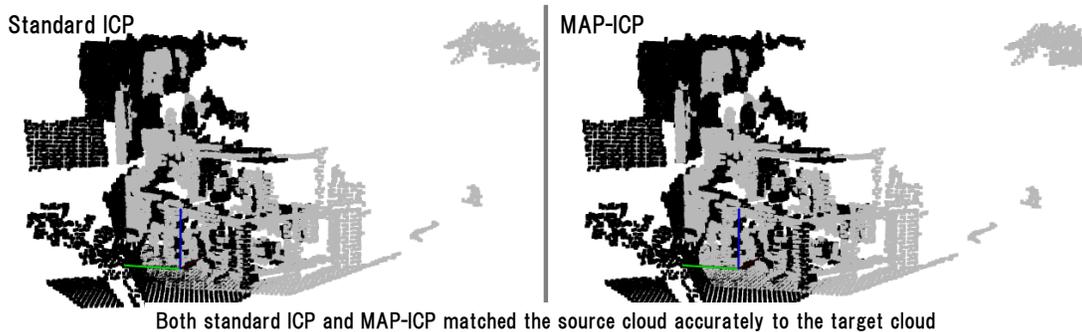


Fig. 8. An example of matching results of standard ICP and MAP-ICP in the room environment

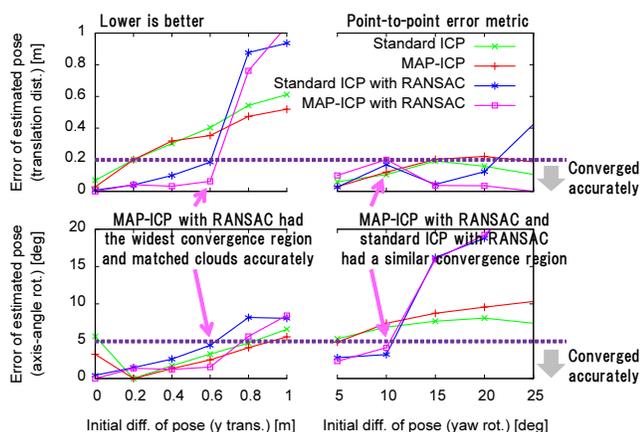


Fig. 9. Errors of estimated pose using point-to-point matching in the room environment

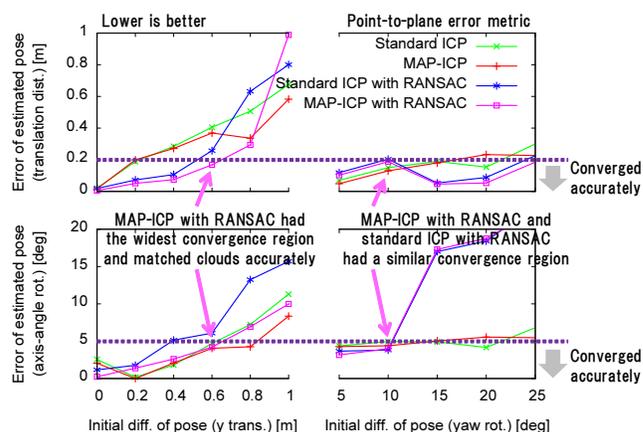


Fig. 10. Errors of estimated pose using point-to-plane matching in the room environment

convergence region over 0.6 m and 10 deg. It matched point clouds accurately even when there was a certain amount of initial pose translation. In the case of initial pose rotation, MAP-ICP with RANSAC and standard ICP with RANSAC matched point clouds with similar accuracy and had a similar convergence region.

Fig. 12 and 13 show the plots of localization error as a function of the initial displacement in the hallway environment. Fig. 12 shows point-to-point matching, while Fig. 13 shows point-to-plane. Similarly to the room environment, MAP-ICP with RANSAC had wide convergence region in both the point-to-point and the point-to-plane cases. In the point-to-point case, the convergence region was almost 0.4 m and 10 deg. In the point-to-plane case, the convergence region was approximately 0.4 m and 15 deg. MAP-ICP with RANSAC matched point clouds accurately even when some amount of initial pose translation was occurred. On the other hand, MAP-ICP (without RANSAC) had wider convergence region than MAP-ICP with RANSAC when a certain amount of initial pose rotation was present. This property was dependent on the shapes of the point clouds. In the point-to-point case especially when there was initial pose rotation for the range of less than 10 deg, MAP-ICP with RANSAC was the most accurate. Then, in the point-to-plane case when there was some amount of initial pose

rotation, MAP-ICP with RANSAC and standard ICP with RANSAC matched point clouds with similar accuracy and had a similar convergence region.

Fig. 15 and 16 show the plots of localization error as a function of the initial displacement in the dynamic environment. Fig. 15 shows point-to-point matching, while Fig. 16 shows point-to-plane. Similarly to the room and the hallway environment, MAP-ICP with RANSAC had the widest convergence region over 0.2 m and 5 deg in both the point-to-point and the point-to-plane cases. It matched point clouds accurately even when a certain amount of initial pose translation was occurred. This was also true in the case where the initial guess was the ground truth, and localization error of MAP-ICP with RANSAC was nearly zero. In the case of initial pose rotation, only MAP-ICP with RANSAC and MAP-ICP matched point clouds accurately.

From the above experimental results, we obtained the conclusions that our proposed method MAP-ICP with RANSAC had the widest convergence region and matched point clouds accurately in both the point-to-point and the point-to-plane cases. In 6DOF matching for 3D shapes, there are a large number of local minima and hence matching is likely to fail. By considering both the likelihood of the measurement and the prior probability of the initial guess with MAP estimation, MAP-ICP will not fall into misalignment and

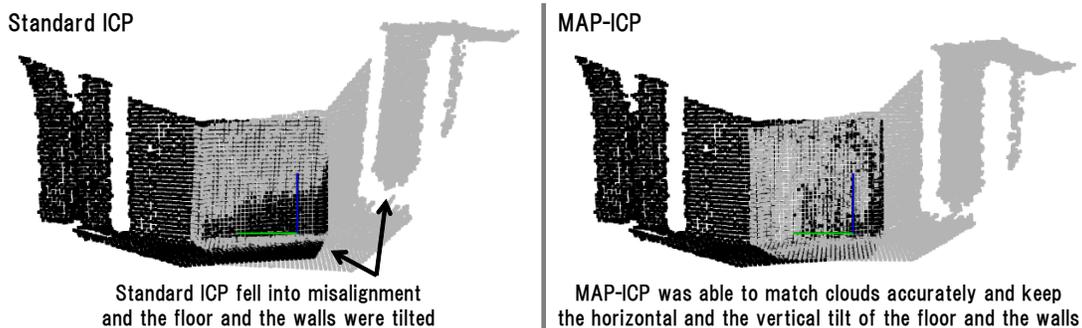


Fig. 11. An example of matching results of standard ICP and MAP-ICP in the hallway environment

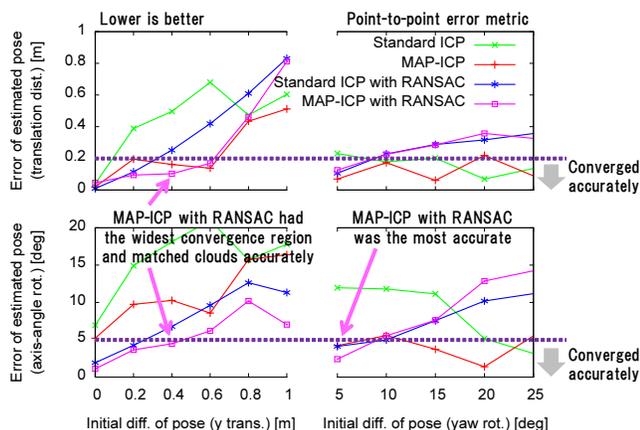


Fig. 12. Errors of estimated pose using point-to-point matching in the hallway environment

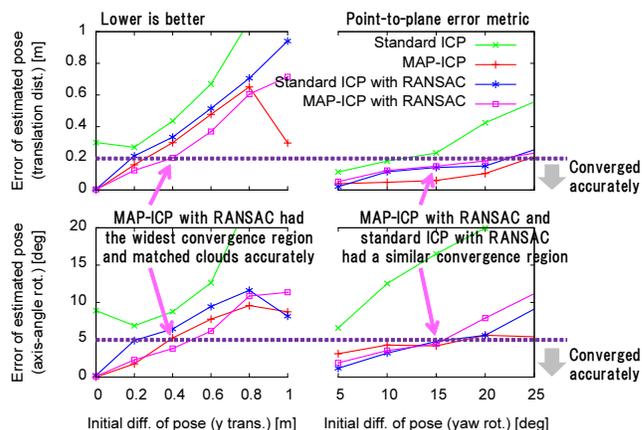


Fig. 13. Errors of estimated pose using point-to-plane matching in the hallway environment

is able to localize robot pose accurately even if we use a depth camera that has large measurement errors and a narrow sensing field of view.

V. CONCLUSION

In this paper, our objective was a new 6DOF matching algorithm for 3D shapes based on probability theory for localization. Since conventional approaches are particularly susceptible to becoming stuck in the numerous local minima, and fall into misalignment, hence we introduced MAP estimation based on Bayes' theorem to the ICP algorithm. Our approach resolves such problems by considering both the likelihood of the measurement and the prior probability of the initial guess for registration in the objective function. We have implemented MAP-ICP algorithm and evaluated localization errors of the method in comparison with conventional methods. The experimental results have shown that our proposed method has the widest convergence region and is able to match point clouds accurately preventing the misalignment problem. Consequently, if we can obtain an appropriate initial guess, it is better to use MAP-ICP rather than the standard ICP for fine registration.

ACKNOWLEDGMENT

This work has partly been supported by the JSPS Grant-in-Aid for JSPS Fellows under contract number 24-2589.

REFERENCES

- [1] Martin Buehler, Karl Iagnemma, and Sanjiv Singh: "The DARPA Urban Challenge", *Springer*, 2010.
- [2] John Markoff: "Google Cars Drive Themselves, in Traffic", *New York Times*, 2010. <http://www.nytimes.com/2010/10/10/science/10google.html>
- [3] Christopher M. Bishop: "Pattern Recognition and Machine Learning", *Springer*, 2006.
- [4] Sebastian Thrun, Wolfram Burgard, and Dieter Fox: "Probabilistic Robotics", *The MIT Press*, 2005.
- [5] Joaquim Salvi, Carles Matabosch, David Fofi, and Josep Forest: "A Review of Recent Range Image Registration Methods with Accuracy Evaluation", *J. of Image and Vision Computing*, Vol. 25, No. 5, pp. 578–596, 2007.
- [6] Paul J. Besl, and Neil D. McKay: "A Method for Registration of 3-D Shapes", *IEEE Trans. on PAMI*, Vol. 14, No. 2, pp. 239–256, 1992.
- [7] Yang Chen, and Gerard Medioni: "Object Modeling by Registration of Multiple Range Images", *J. of Image and Vision Computing*, Vol. 10, No. 3, pp. 145–155, 1992.
- [8] Aleksandr V. Segal, Dirk Haehnel, and Sebastian Thrun: "Generalized-ICP", *Proc. of RSS2009*, 2009.
- [9] Szymon Rusinkiewicz, and Marc Levoy: "Efficient Variants of the ICP Algorithm", *Proc. of 3DIM2001*, 2001.
- [10] Feng Lu, and Evangelos Miliotis: "Robot Pose Estimation in Unknown Environments by Matching 2D Range Scans", *J. of Intelligent and Robotic Systems*, Vol. 18, No. 3, pp. 249–275, 1997.
- [11] Andreas Nuchter, Kai Lingemann, Joachim Hertzberg, and Hartmut Surmann: "6D SLAM - 3D Mapping Outdoor Environments", *J. of Field Robotics*, Vol. 24, No. 8–9, pp. 699–722, 2007.
- [12] Yoshitaka Hara, Akira Oshima, Yukihiko Ono, Azusa Amino, Kenjiro Yamamoto: "Development of Autonomous Navigation Technology Adapted to Crowded Pedestrian Streets and Evaluations in Real World

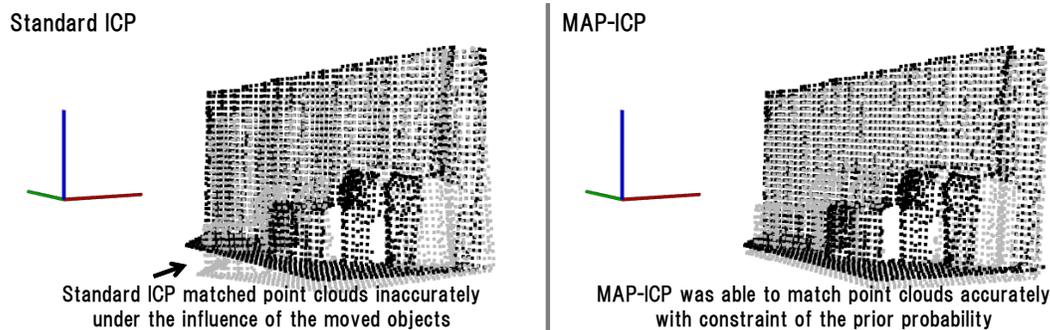


Fig. 14. An example of matching results of standard ICP and MAP-ICP in the dynamic environment

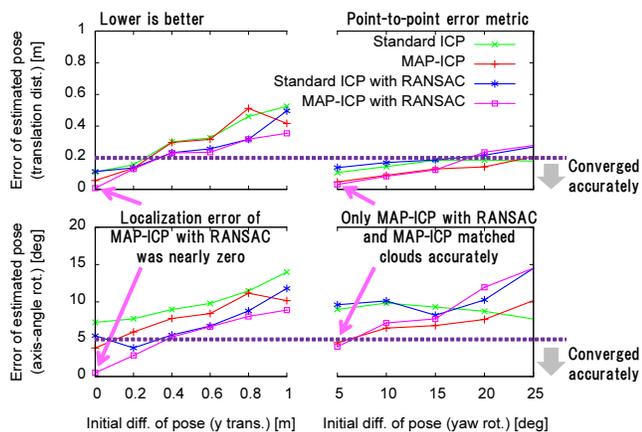


Fig. 15. Errors of estimated pose using point-to-point matching in the dynamic environment

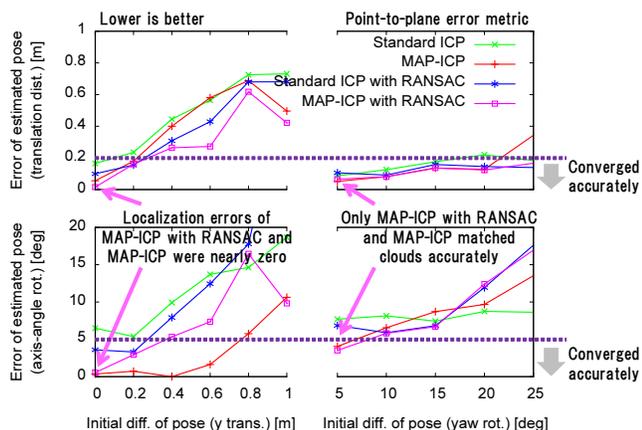


Fig. 16. Errors of estimated pose using point-to-plane matching in the dynamic environment

Using Experimental Robot Sofara-T”, *J. of the Robotics Society of Japan*, Vol. 30, No. 3, pp. 287–295, 2012. (in Japanese)

- [13] Shun’ichi Kaneko, Tomonori Kondo, and Atsushi Miyamoto: “Robust Matching of 3D Contours Using Iterative Closest Point Algorithm Improved by M-Estimation”, *J. of Pattern Recognition*, Vol. 36, No. 9, pp. 2041–2047, 2003.
- [14] Sebastien Granger, and Xavier Pennec: “Multi-Scale EM-ICP: A Fast and Robust Approach for Surface Registration”, *Proc. of ECCV2002*, 2002.
- [15] Roland Siegwart, Illah R. Nourbakhsh, and Davide Scaramuzza: “Introduction to Autonomous Mobile Robots”, *The MIT Press*, 2011.

- [16] Edwin Olson: “Real-Time Correlative Scan Matching”, *Proc. of ICRA2009*, 2009.
- [17] Kurt Konolige, and Ken Chou: “Markov Localization Using Correlation”, *Proc. of IJCAI1999*, 1999.
- [18] Rainer Kummerle, Rudolph Triebel, Patrick Pfaff, and Wolfram Burgard: “Monte Carlo Localization in Outdoor Terrains Using Multi-Level Surface Maps”, *J. of Field Robotics*, Vol. 25, No. 6–7, pp. 346–359, 2008.
- [19] Taro Suzuki, Yoshiharu Amano, and Takumi Hashizume: “6-DOF Localization for a Mobile Robot Using Outdoor 3D Point Clouds”, *J. of Robotics and Mechatronics*, Vol. 22, No. 2, pp. 158–166, 2010.
- [20] Frank Dellaert, Dieter Fox, Wolfram Burgard, and Sebastian Thrun: “Monte Carlo Localization for Mobile Robots”, *Proc. of ICRA1999*, 1999.
- [21] Fletcher Dunn, and Ian Parberry: “3D Math Primer for Graphics and Game Development”, *CRC Press*, 2011.
- [22] Martin A. Fischler, and Robert C. Bolles: “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography”, *Comm. of the ACM*, Vol. 24, No. 6, pp. 381–395, 1981.
- [23] Point Cloud Library.
<http://pointclouds.org/>