2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
Madrid, Spain, October 1-5, 2018

Dynamic Scaling Factors of Covariances for Accurate 3D Normal Distributions Transform Registration

Hyunki Hong and B. H. Lee

Abstract— **Distribution-to-distribution normal distributions transform (NDT-D2D) is one of the fast point set registrations. Since the normal distributions transform (NDT) is a set of normal distributions generated by discrete and regular cells, local minima of the objective function is an issue of NDT-D2D. Also, we found that the objective function based on L_2 distance between distributions has a negative correlation with rotational alignment. To overcome the problems, we present a method using dynamic scaling factors of covariances to improve the accuracy of NDT-D2D. Two scaling factors are defined for the preceding and current NDTs respectively, and they are dynamically varied in each iteration of NDT-D2D. We implemented the proposed method based on conventional NDT-D2D and probabilistic NDT-D2D and compared to the NDT-D2D with fixed scaling factors using KITTI benchmark data set. Also, we experimented estimating odometry with an initial guess as an application of distribution-to-distribution probabilistic NDT (PNDT-D2D) with the proposed method. As a result, the proposed method improves both translational and rotational accuracy of the NDT-D2D and PNDT-D2D.**

I. INTRODUCTION

Point set registration is a critical task required in many fields such as computer vision, medical imaging, graphics, and robotics [1]–[4]. For the simultaneous localization and mapping (SLAM) problem in the robotics field, the odometry of a robot can be estimated by registering a pair of point coordinate sets or feature point sets collected from vision sensors or rangefinders. So far, the lidar odometry and mapping (LOAM) estimates the odometry by the thread registering point clouds and updates the odometry by the mapping thread [5]. In this paper, the proposed registration method is aimed at the odometry estimation.

Normal distributions transform (NDT) is a compact spatial representation which represents the geometric information by the normal distributions (ND) converted from the points, and the registration of the preceding NDT M_{pre} and current point set P_{cur}, also known as point-to-distribution NDT (NDT-P2D), is faster than the variants of iterative closest point (ICP) [6]. The reasons are that the conversion and searching for correspondence are fast due to the generation using regular cells, the number of NDs is much smaller than the number of points. After the methods registering Gaussian mixture model (GMM) has been studied [7], [8], distribution-to-distribution NDT (NDT-D2D) was also proposed [9]. Since the number of NDs of the current NDT M_{cur} is also much smaller than the number of points, its registration rate overwhelms NDT-P2D [9].

In this paper, we propose a registration improved by dynamic scaling factors (DSF). We define two scaling factors of covariances s_{pre} and s_{cur} for M_{pre} and M_{cur} respectively. The proposed registration is based on NDT-D2D, and it can be divided into three steps. First is NDT-P2D-like registration which can register without the negative correlation between L_2 distance and rotational alignment. Second is NDT-D2D with decreasing s_{cur} and s_{pre} to roughly obtain the transformation. Third is NDT-D2D with fixed s_{cur} and s_{pre} determined by the translation vector. We implemented the registration based on conventional NDT-D2D and distribution-to-distribution probabilistic NDT (PNDT-D2D) [10]. Experiments are conducted with those methods and KITTI benchmark data set. As a result, the proposed DSF improved the conventional NDT-D2D and PNDT-D2D.

II. RELATED WORK

A. NDT representations

NDT is a set of NDs representing the spatial information. The conventional NDT subdivides a point set by non-overlapped regular cells and converts the point set $P = \{\hat{\mathbf{p}}_i = (x_i, y_i, z_i)\}, i = 1, 2, ..., n$ in the cell to the normal distribution (ND) $\mathcal{N}(\mu, \Sigma)$ by

$$\mu = \frac{1}{n}\sum_{i=1}^{n}\hat{\mathbf{p}}_i, \tag{1}$$

$$\Sigma = \sum_{i=1}^{n}(\hat{\mathbf{p}}_i - \mu)(\hat{\mathbf{p}}_i - \mu)^T, \tag{2}$$

where μ is the mean vector and Σ is the covariance matrix [11], [12]. In our previous work, we presented a probabilistic NDT (PNDT) representation which computes $\mu^{'}$ and $\Sigma^{'}$ as

$$\mu^{'} = \frac{1}{n}\sum_{i=1}^{n_P}\hat{\mathbf{p}}_i, \tag{3}$$

$$\Sigma^{'} = \frac{1}{n}\sum_{i=1}^{n_P}\Sigma_{p,i} + \Sigma, \tag{4}$$

where Σ_p is the uncertainty of $\mathbf{p}$ [10]. Since PNDT converts points whose number is smaller than 4 in a cell into one distribution based on uncertainty, it can represent the space with relatively small cells. Therefore, to estimate transformation accurately with small cells, we suggest to use PNDT.

* This work was supported in part by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (No. 2017R1A2B2002608) and in part by a grant to Bio-Mimetic Robot Research Center Funded by Defense Acquisition Program Administration, and by Agency for Defense Development (UD130070ID).

Hyunki Hong and B.H. Lee are with the Department of Electrical and Computer Engineering(Automation and Systems Research Institute), Seoul National University, Seoul 08826, Korea {ckowl,bhlee}@snu.ac.kr

B. Registrations using NDT

Given an preceding NDT M_{pre}, the likelihood of an arbitrary point $\mathbf{p}_a$ being observed can be defined as

$$p(\mathbf{p}_a|M_{pre}) = \sum_{i=1}^{n_M} w_i \mathcal{N}(\mathbf{p}_a|\mu_i,\Sigma_i), \tag{5}$$

where w_i is weight and n_M is the number of NDs [9], [12]. Since n_M is usually much smaller than n, NDT can not only represent the space by the less memory, but also rapidly register the point cloud. The objective function of the NDT registration is defined based on (5). Given an preceding NDT M and a current point set P_{cur}, the objective function of NDT-P2D is

$$f(M_{pre},P_{cur},\Theta) = \sum_{i=1}^{|P_{cur}|} -d_1 exp\left(-\frac{d_2}{2} g(\mathbf{p}_i,\mu_{c,i},\Sigma_{c,i},\Theta)\right), \tag{6}$$

$$g(\mathbf{p},\mu_c,\Sigma_c,\Theta) = (T(\mathbf{p},\Theta)-\mu_c)^T \Sigma_c^{-1} (T(\mathbf{p},\Theta)-\mu_c), \tag{7}$$

where Θ is the parameter of transformation, d_1 and d_2 are positive regularizing factors, (μ_c,Σ_c) is the correspondence of $\mathbf{p}$, and $T(\mathbf{p},\Theta)$ is the function transforming $\mathbf{p}$ with Θ.

On the other hand, the objective function of NDT-D2D is the modification of L_2 distance between two NDTs [9]. Given the preceding and current NDTs M_{pre} and M_{cur}, the L_2 distance between two NDTs can be derived as

$$f(M_{pre},M_{cur},\Theta) = \sum_{i=1}^{n_{cur}} \sum_{j=1}^{n_{pre}} w_{ij} \mathcal{N}(0|\mu_{ij},\Sigma_{ij}), \tag{8}$$

$$\mu_{ij} = T(\mu_i,\Theta) - \mu_j, \tag{9}$$

$$\Sigma_{ij} = R^T \Sigma_i R + \Sigma_j, \tag{10}$$

where R is the rotation matrix determined by Θ, n_{pre} and n_{cur} are the numbers of NDs in M_{pre} and M_{cur}, respectively, (μ_i,Σ_i) and (μ_j,Σ_j) are the parameter sets of the ith normal distribution in M_{cur} and the jth normal distribution in M_{pre}, respectively, and the objective function is defined as

$$f(M_{pre},M_{cur},\Theta) = \sum_{i=1}^{n_{cur}} \sum_{j=1}^{n_{pre}} d_{ij}, \tag{11}$$

$$d_{ij} = -d_1 exp\left(-\frac{d_2}{2} g_{ij}\right), \tag{12}$$

$$g_{ij} = \mu_{ij}^T \Sigma_{ij}^{-1} \mu_{ij}. \tag{13}$$

Since n_{cur} is much smaller than $|P_{cur}|$, NDT-D2D processes at higher rate than NDT-P2D.

C. Background

So far, various approaches have been proposed to improve the performance of NDT registration, and we found a common property which is smoothing objective function to reduce the local minima problem. This property can be seen in following approaches.

Bilinear 2D NDT and trilinear 3D NDT can solve the problem of discreteness by generating overlapped cells [11], [13]. Even though the trilinear NDT-D2D shows high accuracy, its computation time is eight or more times longer than the conventional NDT-D2D. The hierarchical approach is another good example. It registers P_{cur} to coarse-to-fine NDTs [12], [14], [15]. Also, a method of choosing resolution by comparing the number of utilized points was proposed to accelerate the NDT-P2D [16]. Recently, an uncertainty approach for GMM is proposed [17]. According to the paper, the NDs can be expanded by the surface and normal uncertainties. Also, the probabilistic approach can expand NDs by the expected mean and covariance considering probability density function (pdf) based on the sensor model [10]. Although these approaches can improve the accuracy of NDT and GMM, we still have to determine the regularizing factors for them experimentally.

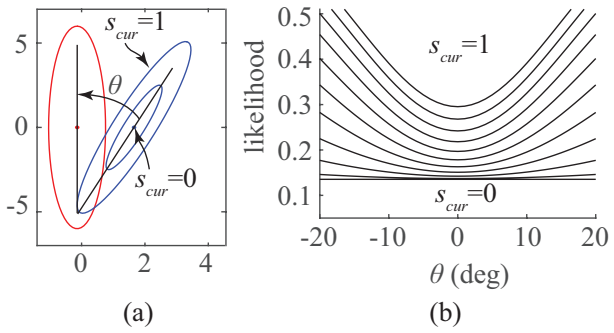

Fig. 1: Negative correlation between L_2 likelihood and rotational alignment. (a) red: 1σ of the preceding covariance, blue: 1σ of the current covariances rotated with θ and scaled with s_{cur}. (b) L_2 likelihood between two distribtions against $|\theta|$ with different s_{cur}.

NDT-D2D has another issue caused by the objective function. Generally, L_2 likelihood is expected to increase as the point sets are aligned. However, it is negatively correlated to the rotational alignment at a location. Given M_{pre} in red and M_{cur} rotated with θ in blue as shown in Fig. 1(a), we can have the graph of L_2 likelihood against θ in Fig. 1(b). As a result, although it is rotationally aligned as $|\theta|$ decreases, the L_2 likelihood decreases as shown in Fig. 1(b).

To overcome the problem, we suggest to initialize s_{cur} to 0 and gradually increase it. In Fig. 1(b), the L_2 likelihood against θ with different s_{cur} from 1 to 0 can be seen. As s_{cur} decreases, the changes caused by θ decreases, and if $s_{cur} = 0$, it is degenerated as NDT-P2D of M_{pre} and mean vectors of M_{cur}.

III. THE PROPOSED METHOD

A. Scaled NDT representation and L_2 distance

NDT-D2D with dynamic scaling factors (NDT-D2D-DSF) is based on the scaled NDT, whose covariance matrices are scaled by a positive scaling factor s. We define a scaled covariance Σ_s as

$$\Sigma_s = s\Sigma. \tag{14}$$

It is equivalent to scale the eigenvalue of Σ with s. The eigenvector matrix V and eigenvalue matrix D of Σ obtained

by eigenvalue decomposition can be expressed as

$$\Sigma = VDV^T, \tag{15}$$

$$D = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}, \tag{16}$$

where σ_i^2 is the variance on ith eigenvector, and the scaled covariance Σ_s can be derived as

$$\Sigma_s = s\Sigma = V(sD)V^T = VD_sV^T. \tag{17}$$

Therefore, as s increases, the variance is swelled and the NDT is smoothed as shown in Fig. 2.

To implement the solution we mentioned in the motivation section, we define two individual scaling factors s_{cur} and s_{pre} for M_{pre} and M_{cur}, respectively, and we modify (10) to

$$\Sigma_{s,ij} = s_{cur}R\Sigma_i R^T + s_{pre}\Sigma_j, \tag{18}$$

and d_2 can be seen as the special case of $s_{cur} = s_{pre} = 1/d_2$. Thus, the proposed objective function can be expressed as

$$f_s(M_{pre}, M_{cur}, \Theta) = \sum_{i=1}^{n_{cur}} \sum_{j=1}^{n_{pre}} d_{s,ij}, \tag{19}$$

$$d_{s,ij} = -d_1 exp\Big(-\frac{1}{2}g_{s,ij}\Big), \tag{20}$$

$$g_{s,ij} = \mu_{ij}^T \Sigma_{s,ij}^{-1} \mu_{ij}. \tag{21}$$

B. NDT-D2D with dynamic scaling factors of covariances

The key idea of the proposed method is varying scaling factors s_{pre} and s_{cur} in each iteration to improve the accuracy of the NDT-D2D. To decrease and increase the scaling factors, we define a function as follows:

$$s(k) = S(k, k_i, k_f, s_i, s_f), \tag{22}$$

where k_i and k_f are the initial and final iterations, respectively, and s_i and s_f are the initial and final scales, respectively.

At the first, since M_{cur} is not rotationally aligned to M_{pre}, the negative correlation problem might exist. Therefore, the strategy is to initialize s_{cur} to 0 so that it becomes the NDT-P2D between M_{pre} and mean vectors of M_{cur}. Also, to attract those mean vectors of M_{cur}, M_{pre} is swelled with $s_{pre} = s_{max}$. After this initialization, the registration performs the following processes in order: scaling NDTs to avoid the negative correlation between L_2 distance and rotational alignment, shrinking NDTs with the same scaling factors, and scaling NDTs considering the translation.

In the first process, since M_{cur} can be expected to be rotationally aligned to M_{pre} as the registration iterates, s_{pre} can gradually decrease from s_{max} to s_1 shrink M_{pre} to regain the shape by $S(k, 0, k_1, s_{max}, s_1)$ and s_{cur} can gradually increase from 0 to s_1 show the shape of M_{cur} by $S(k, 0, k_1, 0, s_1)$). The process continues until $s_{cur} = s_{pre}$ in k_1th iteration as shown in Fig. 3.

In the second process, s_{cur} and s_{pre} decrease as the same function $S(k, k_1, k_2, s_1, s_{tran}(k))$ until $s_{new} = s_{tran}$ in

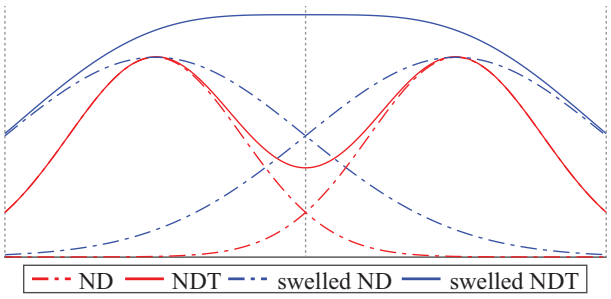

Fig. 2: Illustration of the scaled NDT.

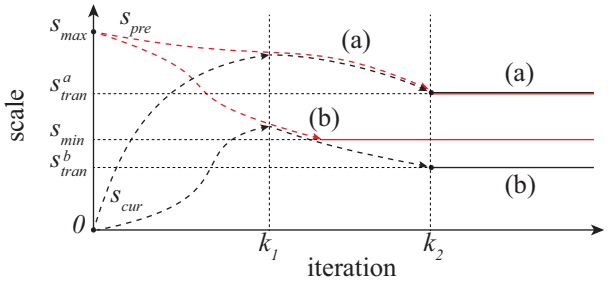

Fig. 3: Illustration of scaling factors varying with iteration.

k_2th iteration. Since $s_{cur} = s_{pre}$, Σ_i and Σ_j have the same influence on the cost, gradient, and Hessian matrix of the objective function. Also, as the scaling factors decrease to s_{tran}, M_{pre} and M_{cur} are shrunken to regain their shapes so that the registration can be expected to estimate the more accurate transformation. If s_{tran} is greater than s_{min}, s_{pre} would be equal to s_{new}. However, since the local minima of the objective function can appear again due to the small s_{tran}, we ask s_{pre} to be at least s_{min} to avoid the local minima.

In the final process, M_{cur} registers to M_{pre} with fixed $s_{cur} = s_{tran}$ and s_{pre}. Iif $s_{tran} > s_{min}$, s_{pre} is equal to s_{tran} as the case (a) in Fig. 3, otherwise s_{pre} is equal to s_{min} as the case (b) in Fig. 3. The process terminates if the condition of the termination is satisfied.

The properties of the processes can be seen in the optimization process. In this paper, we choose Newton method as an example. To estimate $\hat{\Theta}$ by minimizing (19), it updates $\hat{\Theta}$ iteratively as

$$\hat{\Theta}_k = \hat{\Theta}_{k-1} + \Delta\hat{\Theta}_k, \tag{23}$$

$$\Delta\hat{\Theta}_k = -\gamma\left(\frac{\partial^2 f_s}{\partial\theta_a \partial\theta_b}\right)^{-1}\left(\frac{\partial f_s}{\partial\theta_a}\right)^T, \tag{24}$$

where γ is an update rate. The gradient and Hessian matrix of (19) are derived in Appendix.

At the initial state which $s_{pre} = s_{max}$ and $s_{cur} = 0$, we have the gradient and Hessian of g_{ij} as follows:

$$\frac{\partial g_{s,ij}}{\partial\theta_a} = \frac{2}{s_{pre}}\mu_{ij}^T\Sigma_i^{-1}\frac{\partial\mu_{ij}}{\partial\theta_a}, \tag{25}$$

$$\frac{\partial^2 g_{s,ij}}{\partial\theta_a\partial\theta_b} = \frac{2}{s_{pre}}\left(\mu_{ij}^T\Sigma_i^{-1}\frac{\partial^2\mu_{ij}}{\partial\theta_a\partial\theta_b} + \frac{\partial\mu_{ij}^T}{\partial\theta_a}\Sigma_i^{-T}\frac{\partial\mu_{ij}}{\partial\theta_b}\right). \tag{26}$$

(24) can be derived as

$$\Delta\hat{\Theta}_t = -\gamma\Big(\sum_{i,j} d_{s,ij}\big(g_{ij}^{(2)} - \frac{1}{s_{pre}} g_{ij}^{(1)^T} g_{ij}^{(1)}\big)\Big)^{-1} \cdot \Big(\sum_{i,j} d_{s,ij} g_{ij}^{(1)}\Big)^T, \tag{27}$$

where g_{ij} is (13). This equation leads to the same update vector obtained by NDT-P2D of M_{pre} scaled by s_{pre} and the mean vectors of M_{cur}. As s_{pre} decreases, the influence of $g_{ij}^{(2)}$ decreases, and d_{ij} also decreases. If s_{pre} is so large that all of d_{ij} are approximately 1, $\Delta\hat{\Theta}_k$ is

$$\Delta\hat{\Theta}_t = -\gamma\Big(\sum_{i,j} g_{ij}^{(2)}\Big)^{-1} \cdot \Big(\sum_{i,j} g_{ij}^{(1)}\Big)^T. \tag{28}$$

It is the upper bound of $\Delta\hat{\Theta}_k$ against s_{pre}, and it is exactly the same update vector obtained by NDT-P2D based on Mahalanobis distance [15].

In the first process, the first and second derivatives of $\Sigma_{s,ij}^{-1}$ increases as s_{cur} increases. It means that the registration not only aligns the mean vectors of M_{cur} to M_{pre} but also aligns the shapes of NDs. In the second and third processes, as s_{pre} and s_{cur} decrease, $d_{s,ij}$ gap increases. Since $d_{s,ij}$s in (39) and (42) act like weights, the pair of NDs having larger L_2 likelihood has more influence on computing update vector.

C. Range of scaling factors

Since distributions of the conventional NDT are constructed by discrete cells, local minima of the objective function (11) exist. Therefore, smoothing the NDT is critical for NDT registration. In this paper, we suggest a method of smoothing linear and planar distributions and a method of smoothing the distributions based on the translation distance.

Assume a dense point set in a cell whose size is l and range is (l_0, l_1) in one dimension, the mean μ_1 is equal to $l_0 + l/2$, and the covariance Σ_1 is equal to $l^2/12$. We consider another distribution which is $\mathcal{N}(\mu_2, \Sigma_2)$, where $\mu_2 = l_0 + 2l$ and $\Sigma_2 = \Sigma_1$, and have a scaled Gaussian mixture function as

$$p(x) = \sum_{i=0}^{1} \exp -\frac{(x-\mu_i)^2}{2s\Sigma_i}, \tag{29}$$

and a gradient of the function expected to be zero is

$$\frac{dp(x)}{dx} = \sum_{i=0}^{1} \frac{(x-\mu_i)}{s\Sigma_i} \exp -\frac{(x-\mu_i)^2}{2s\Sigma_i}. \tag{30}$$

By parameterizing x as $x(t) = \mu_1 + tl, t \in (0,1)$, we have the scaling factor $s(t)$ making the gradient zero as

$$s(t) = \frac{12(2t-1)}{\log(1/t-1)}. \tag{31}$$

By (31), it is possible to determine a scaling factor to set the gradient at x to 0. In this paper, we define a reference scale s_{min} which sets the gradient to zero at the middle of two means, and it is obtained by $\lim_{t\to 0.5} s(t) = 3$ which smooths distributions as shown in Fig. 2.

We also suggest a method to determine the target scaling factor s_{tgt} based on translation δ. As shown in Fig. 4, the

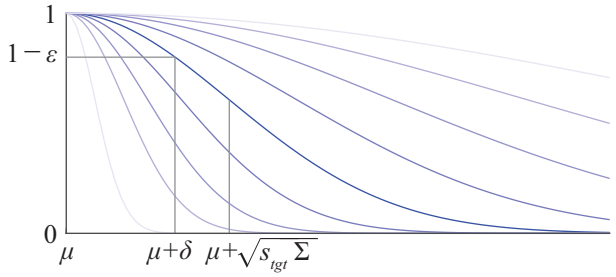

Fig. 4: Illustration of determining s_{tgt} based on the likelihood difference limit ε at $\mu + \delta$.

s_{tgt} is defined as the factor scaling the distribution to have likelihood equal to $1-\varepsilon$ at $\mu + \delta$ so that

$$1 - \varepsilon = exp\left(-\frac{\delta^2}{2s_{max}\Sigma}\right). \tag{32}$$

We apply the same assumption $\Sigma = l^2/12$ as mentioned before, and s_{tgt} can be computed by

$$s_{tgt} = -\frac{6\delta^2}{l^2 log(1-\varepsilon)}. \tag{33}$$

To compute s_{max} in section III-B, we substitute the maximum velocity v_{max} as δ. Also, the translation $|\mathbf{t}|$ can be substituted into (33) to obtain s_{tran} in section III-B.

IV. EXPERIMENTS

To verify the performance of the registration with the dynamic scaling factor, first, we conducted experiments on the NDT-D2D and PNDT-D2D without initial guess in the following cases: different scaling factors, different s_{cur} with s_{pre} increasing from 0 to s_{ref}, and the proposed method. Second, we experimented with estimating odometry by the conventional PNDT-D2D and PNDT-D2D with the proposed method as an application. In this case, the registration has a previous transformation as the initial guess.

A. Dataset and evaluation

Since KITTI data set collected by 64 channel lidar Velodyne mounted on a car provides accurate ground truths [18], we choose the data set as a benchmark to compare the accuracy of the paramter vector of the transformation $\hat{\Theta} = (t_x, t_y, t_z, a_x, a_y, a_z)$, where t_x, t_y, t_z are translations in Euclidean space, a_x, a_y, a_z are the XYZ-Euler angles. For each case, we experimented with ten sequences, total 23190 pairs.

For the first experiment, it performed the simple scan-to-scan registration without any initial guesses. Therefore, we computed the ground truth of the pose variation ΔT which can be expressed as a homogeneous matrix:

$$\Delta T = \begin{bmatrix} \Delta R & \Delta\mathbf{t} \\ 0 & 1 \end{bmatrix}, \tag{34}$$

where ΔR is the rotational variation and $\Delta\mathbf{t}$ is the translational variation, and we compared the estimated transformation matrix $\Delta\hat{T}$ to compute the error E as

$$E = \Delta\hat{T}^{-1}\Delta T = \begin{bmatrix} E_R & E_t \\ 0 & 1 \end{bmatrix}. \tag{35}$$

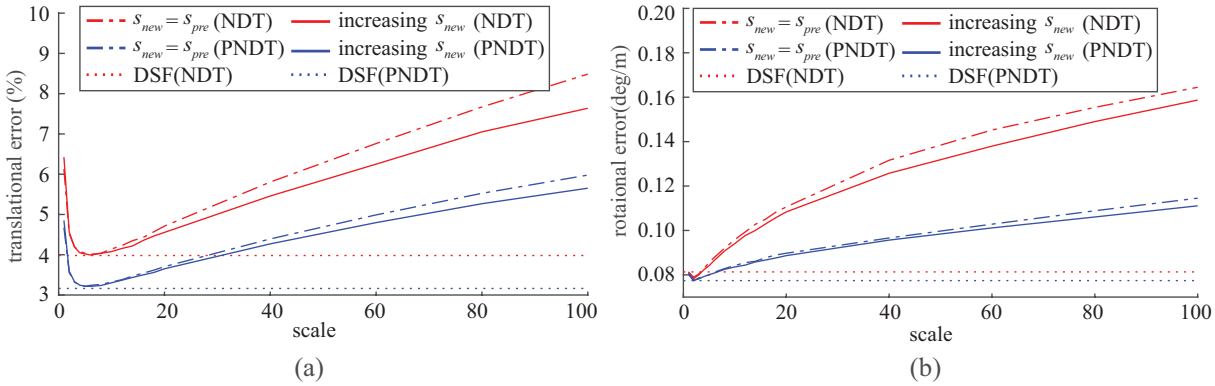

Fig. 5: Median of the translational and rotational errors against the scale. The lowest translational error is found at $s = 6$, and the lowest rotational error is found at $s = 2$. (a) translational errors (b) rotational errors.

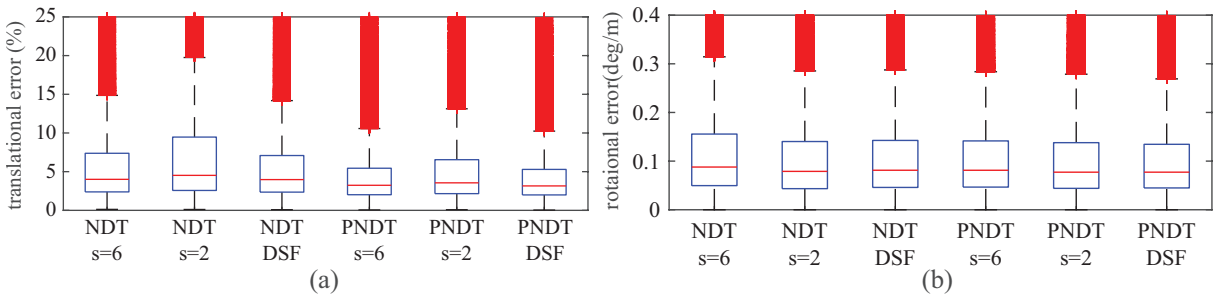

Fig. 6: Errors of NDT-D2D and PNDT-D2D with $s = 6$, $s = 2$, and DSF, respectively. (a) box plots of the translational errors (b) box plots of the rotational errors.

In addition, the translational error is computed as

$$e_t = \frac{|E_t|}{|\Delta t|} \times 100\%, \tag{36}$$

and the rotational error is computed as

$$e_r = \frac{1}{|\Delta t|} cos^{-1}\left(\frac{1}{2}\left(Tr(E_R) - 1\right)\right). \tag{37}$$

For the second experiment, we followed the KITTI's method which evaluates the errors every second with lengths (100,200,...,800) meters by (36) and (37).

B. Implementation

To accelerate the implementation, we upgraded the implementation in our previous work [10]. We combined generating and registering processes into one thread and constructed a thread for each son node of the root of octree [19]. Each thread individually divides input point clouds, computes the mean and covariance in each cell, searching for the correspondence, and computes the score, gradient, and Hessian of the correspondence, and the root node collects the outputs of eight threads. For the parameters, the cell size is set to 4m, the maximum iteration number is set to 40, and the optimization terminates if $|\Delta\hat{\Theta}|$, the norm of the update vector, is less than 10^{-6}.

C. Evaluation of the proposed method without initial guess

For the first case with fixed scaling factors, the NDT-D2D and PNDT-D2D register point sets with $s_{pre} = s_{cur} \in S = \{1,2,3,4,5,6,8,10,12,14,16,18,20,40,60,80,100\}$. For the second case with dynamic s_{cur} and fixed s_{pre}, the NDT-D2D and PNDT-D2D perform with $s_{pre} \in S$ and s_{cur} increasing from 0 to s_{pre}. For the third case of the proposed DSF, we used the linear function as an example of $S(k, k_i, k_f, s_i, s_f)$:

$$s(k) = S_{de}(k, k_1, s_1, s_2) = \frac{s_f - s_i}{k_f - k_i} k, \tag{38}$$

and we set k_1 to 4 and k_2 to 7. To compute s_{max}, we assumed that the maximum velocity of the car is 180km/h which is equal to 5m/Hz, and we substituted 5 as δ into (33). As a result, we obtained the median of the translational and rotational errors as shown in Fig. 5.

In Fig. 5, the lowest translational error of NDT-D2D can be found at $s = 6$, and the lowest rotational error can be found at $s = 2$. Also, the lowest translational and rotational errors of PNDT-D2D can be found at $s = 6$ and $s = 2$ respectively. Also, the second case which has s_{cur} increasing from 0 to s_{pre} shows the lower errors than the first case for both NDT-D2D and PNDT-D2D.

To compare the third case to the best results, we extract the results of $s = 2$ and $s = 6$ and demonstrate them in the form of boxplots as shown in Fig. 6. The translational and rotational errors of NDT-D2D-DSF and PNDT-D2D with DSF (PNDT-D2D-DSF) show the lower medians than the case $s = 6$ and $s = 2$ respectively. Also, the interquartile range (IQR) of NDT-D2D-DSF is narrower than NDT-D2D with $s = 6$ in Fig. 6(a). On the other hand, since it is difficult to compare

TABLE I: Accuracy of odometry estimated by PNDT-D2D and PNDT-D2D-DSF.

seq.	translational error(%)								rotational error (10^{-3}deg/m)							
	cell size=0.5m		cell size=1m		cell size=2m		cell size=4m		cell size=0.5m		cell size=1m		cell size=2m		cell size=4m	
	PNDT	DSF	PNDT	DSF	PNDT	DSF	PNDT	DSF	PNDT	DSF	PNDT	DSF	PNDT	DSF	PNDT	DSF
0	2.11	1.58	2.31	1.79	1.72	1.63	2.00	1.93	11.3	10.5	11.5	11.4	12.2	12.3	15.6	15.4
1	3.06	2.21	3.22	2.39	3.46	2.17	6.91	2.38	24.5	14.6	16.4	12.0	14.9	13.2	14.2	13.2
2	10.13	2.12	3.14	2.02	4.13	2.56	7.48	5.89	66.8	11.2	13.4	12.4	14.1	13.8	26.6	24.3
3	2.78	2.10	3.64	2.85	3.22	1.79	4.39	4.50	9.1	8.3	9.2	7.8	11.4	11.3	12.7	11.9
4	80.50	91.22	3.42	2.29	4.96	2.82	14.83	3.69	6.5	3.2	8.8	8.6	7.8	7.7	8.3	10.0
5	1.59	1.53	1.72	1.63	1.66	1.64	1.83	1.82	12.5	10.7	10.9	10.7	11.9	12.0	13.7	14.0
6	12.67	1.39	2.06	1.50	1.99	1.77	1.70	1.68	13.6	9.4	12.9	13.2	11.5	11.3	13.2	13.2
7	0.80	0.92	0.94	1.07	1.22	1.03	1.62	1.70	6.7	4.2	5.6	5.9	6.7	5.9	12.8	13.4
8	1.92	2.07	2.12	2.16	2.04	4.64	2.46	2.49	11.9	11.3	11.0	11.5	12.7	18.1	15.6	16.9
9	1.43	1.44	1.73	1.76	2.59	2.20	6.10	4.19	15.1	13.5	15.2	15.5	17.8	18.2	34.7	30.3
10	2.07	2.01	2.52	2.38	2.55	2.00	3.22	3.23	12.5	12.7	12.5	10.5	17.1	15.4	20.8	20.3
mean	10.82	9.87	2.44	1.99	2.69	2.20	4.78	3.05	17.3	10.0	11.6	10.9	12.6	12.7	17.1	16.6

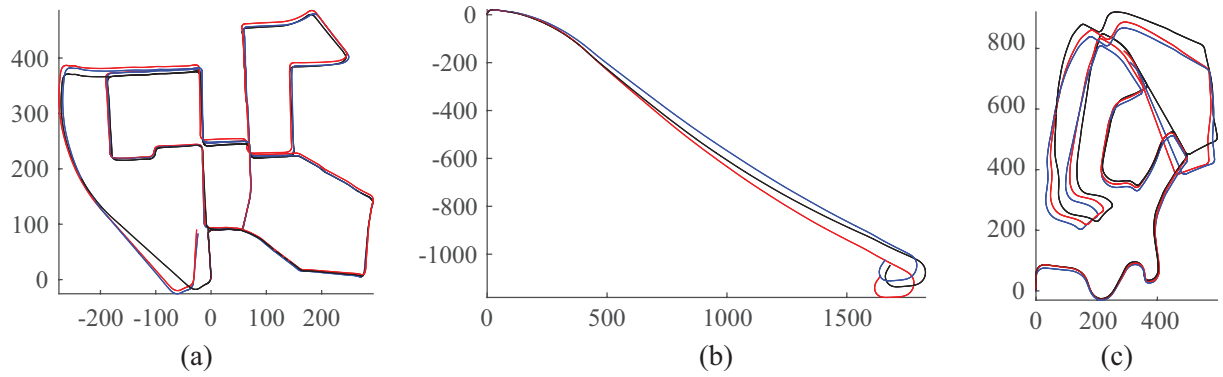

Fig. 7: Results of odometry estimation with cell size l=1m. Black odometry is ground truth, blue one is estimated by PNDT-D2D, and red one is estimated by PNDT-D2D-DSF. (a) 0th sequence. (b) 1st sequence. (c) 2nd sequence.

PNDT-D2D-DSF to PNDT-D2D with the fixed scaling factor, we have values to show the improvements as follows. For the translational error, the IQR of PNDT-D2D-DSF 3.2982 is similar to the IQR of PNDT-D2D with $s = 6$ 3.4233, and the median of PNDT-D2D-DSF 3.1573 is less than the median of PNDT-D2D with $s = 6$ 3.239 about 2.52%. Also, for the rotational error, the IQR of PNDT-D2D-DSF 0.0897 is also similar to the IQR of PNDT-D2D with $s = 6$ 0.0949, and the median of PNDT-D2D-DSF 0.0039 is less than the median of PNDT-D2D with $s = 2$ 0.0480 about 4.8%.

D. Application of odometry estimation

The main purposes of this experiment are not only the application of the proposed method but also the adaptivity to the different cell sizes. We expected that s_{tran}, the scaling factors considering translation in the second and third process, would increase as cell size decreases since δ is similar in (33) while l is decreasing. Hence, we conducted the experiment with cell size l equal to 0.5m, 1m, 2m, and 4m. Since PNDT-D2D-DSF shows better accuracy than NDT-D2D-DSF, we chose PNDT-D2D-DSF to compare to the conventional PNDT-D2D. To improve the accuracy, PNDT-D2D and PNDT-D2D-DSF have the previous transformation as the initial guess. As a result of 0th to 10th sequences, the errors against the cell size are as shown in Table I. The averages of the translational and rotational errors of PNDT-D2D-DSF are almost smaller than those of PNDT-D2D. Also, the estimated odometry of PNDT-D2D-DSF demonstrated in Fig. 7 are closer to the ground truth than those of PNDT-D2D. However, the errors from 7th to 9th sequences are higher than those of PNDT-D2D. The reason in the 7th sequence is that the estimated velocity becomes small on the street like a corridor. Since DSF smoothes the objective function, the transformation can be converged into 0, which leads to the lower cost than the ground truth. Also, the reason in the 8th sequence is that the environment at the first is so extensive that the point cloud forms a shape of a disk. Hence, it leads to the cost estimating 0 is lower than the ground truth.

V. CONCLUSION AND FUTURE WORK

We proposed a dynamic scaling factor approach to improve the accuracy of NDT registration. To avoid the negative correlation between L_2 distance and rotational alignment, the proposed method DSF initially sets s_{cur} to 0. Also, to smooth the objective function and to avoid the discreteness of NDs, DSF sets the range for s_{pre}.

We conducted two experiments. First, we compared the registration with the proposed DSF to the registrations with different fixed scaling factors and the registrations with s_{cur} increasing from 0. As a result, the proposed DSF can improve the accuracy of NDT-D2D and PNDT-D2D. Second, we

experimented with estimating odometry as an application of PNDT-D2D with the proposed method. As a result, the accuracy of PNDT-D2D-DSF is higher than the conventional PNDT-D2D. However, we also found a limitation of the proposed DSF.

To overcome the limitation, in the future we will extend this approach to scale on particular directions on eigenvectors of covariances, and we will extend the approach to the hierarchical NDT-D2D to estimate transformation more accurately at the initial state.

VI. APPENDIX

The gradient vector of (19) can be derived as follows:

$$\frac{\partial f_s(M_{pre}, M_{cur}, \Theta)}{\partial \theta_a} = \sum_{i=1}^{n_{cur}} \sum_{j=1}^{n_{pre}} -\frac{1}{2} d_{s,ij} \frac{\partial g_{s,ij}}{\partial \theta_a}, \tag{39}$$

$$\frac{\partial g_{s,ij}}{\partial \theta_a} = 2\mu_{ij}^T \Sigma_{s,ij}^{-1} \frac{\partial \mu_{ij}}{\partial \theta_a} + \mu_{ij}^T \frac{\partial (\Sigma_{s,ij}^{-1})}{\partial \theta_a} \mu_{ij}, \tag{40}$$

$$\frac{\partial (\Sigma_{s,ij}^{-1})}{\partial \theta_a} = -s_{cur} \Sigma_{s,ij}^{-1} \left(R\Sigma_i \frac{\partial R^T}{\partial \theta_a} + \frac{\partial R}{\partial \theta_a} \Sigma_i R^T \right) \Sigma_{s,ij}^{-1}. \tag{41}$$

Also, the Hessian matrix can be derived as follows:

$$\frac{\partial^2 f_s(M_{pre}, M_{cur}, \Theta)}{\partial \theta_a \partial \theta_b} = \sum_{i=1}^{n_{cur}} \sum_{j=1}^{n_{pre}} \frac{\partial^2 d_{s,ij}}{\partial \theta_a \partial \theta_b}, \tag{42}$$

$$\frac{\partial^2 d_{s,ij}}{\partial \theta_a \partial \theta_b} = -\frac{1}{2} d_{s,ij} \left(\frac{\partial^2 g_{s,ij}}{\partial \theta_a \partial \theta_b} - \frac{1}{2} \frac{\partial g_{s,ij}^T}{\partial \theta_a} \frac{\partial g_{s,ij}}{\partial \theta_b} \right), \tag{43}$$

$$\begin{aligned} \frac{\partial^2 g_{s,ij}}{\partial \theta_a \partial \theta_b} &= 2\mu_{ij}^T \Sigma_{s,ij}^{-1} \frac{\partial^2 \mu_{ij}}{\partial \theta_a \partial \theta_b} + 2 \frac{\partial \mu_{ij}}{\partial \theta_a}^T \Sigma_{s,ij}^{-1} \frac{\partial \mu_{ij}}{\partial \theta_b} \\ &+ 2\mu_{ij}^T \frac{\partial (\Sigma_{s,ij}^{-1})}{\partial \theta_b} \frac{\partial \mu_{ij}}{\partial \theta_a} + 2\mu_{ij}^T \frac{\partial (\Sigma_{s,ij}^{-1})}{\partial \theta_a} \frac{\partial \mu_{ij}}{\partial \theta_b} \\ &+ \mu_{ij}^T \frac{\partial^2 (\Sigma_{s,ij}^{-1})}{\partial \theta_a \partial \theta_b} \mu_{ij}, \end{aligned} \tag{44}$$

$$\begin{aligned} \frac{\partial^2 (\Sigma_{s,ij}^{-1})}{\partial \theta_a \partial \theta_b} &= -s_{cur} \Sigma_{s,ij}^{-1} \left(R\Sigma_i \frac{\partial R}{\partial \theta_a}^T + \frac{\partial R}{\partial \theta_a} \Sigma_i R^T \right) \frac{\partial (\Sigma_{s,ij}^{-1})}{\partial \theta_b} \\ &- s_{cur} \frac{\partial (\Sigma_{s,ij}^{-1})}{\partial \theta_b} \left(R\Sigma_i \frac{\partial R}{\partial \theta_a}^T + \frac{\partial R}{\partial \theta_a} \Sigma_i R^T \right) \Sigma_{s,ij}^{-1} \\ &- s_{cur} \Sigma_{s,ij}^{-1} \left(\frac{\partial R}{\partial \theta_a} \Sigma_i \frac{\partial R^T}{\partial \theta_b} + \frac{\partial^2 R}{\partial \theta_a \partial \theta_b} \Sigma_i R^T \right. \\ &\left. + \frac{\partial R}{\partial \theta_b} \Sigma_i \frac{\partial R^T}{\partial \theta_a} + R\Sigma_i \frac{\partial^2 R^T}{\partial \theta_a \partial \theta_b} \right) \Sigma_{s,ij}^{-1}. \end{aligned} \tag{45}$$

REFERENCES

[1] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[2] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE transactions on medical imaging*, vol. 32, no. 7, pp. 1153–1190, 2013.

[3] G. K. Tam, Z.-Q. Cheng, Y.-K. Lai, F. C. Langbein, Y. Liu, D. Marshall, R. R. Martin, X.-F. Sun, and P. L. Rosin, "Registration of 3d point clouds and meshes: a survey from rigid to nonrigid," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 7, pp. 1199–1217, 2013.

[4] F. Pomerleau, F. Colas, R. Siegwart, *et al.*, "A review of point cloud registration algorithms for mobile robotics," *Foundations and Trends® in Robotics*, vol. 4, no. 1, pp. 1–104, 2015.

[5] J. Zhang and S. Singh, "Low-drift and real-time lidar odometry and mapping," *Autonomous Robots*, vol. 41, no. 2, pp. 401–416, 2017.

[6] M. Magnusson, A. Nuchter, C. Lorken, A. J. Lilienthal, and J. Hertzberg, "Evaluation of 3d registration reliability and speed-a comparison of icp and ndt," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 3907–3912.

[7] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp." in *Robotics: science and systems (RSS)*, vol. 2, no. 4, 2009.

[8] B. Jian and B. C. Vemuri, "Robust point set registration using gaussian mixture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1633–1645, 2011.

[9] T. Stoyanov, M. Magnusson, H. Andreasson, and A. J. Lilienthal, "Fast and accurate scan registration through minimization of the distance between compact 3d ndt representations," *The International Journal of Robotics Research (IJRR)*, vol. 31, no. 12, pp. 1377–1393, 2012.

[10] H. Hong and B. H. Lee, "Probabilistic normal distributions transform representation for accurate 3d point cloud registration," in *Press IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*. IEEE, 2017.

[11] P. Biber and W. Straßer, "The normal distributions transform: A new approach to laser scan matching," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, vol. 3. IEEE, 2003, pp. 2743–2748.

[12] M. Magnusson, "The three-dimensional normal-distributions transform: an efficient representation for registration, surface analysis, and loop detection," Ph.D. dissertation, Örebro universitet, 2009.

[13] M. Magnusson, A. Nuchter, C. Lorken, A. J. Lilienthal, and J. Hertzberg, "Evaluation of 3d registration reliability and speed-a comparison of icp and ndt," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*. IEEE, 2009, pp. 3907–3912.

[14] E. Takeuchi and T. Tsubouchi, "A 3-d scan matching using improved 3-d normal distributions transform for mobile robotic mapping," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*. IEEE, 2006, pp. 3068–3073.

[15] C. Ulaş and H. Temeltaş, "3d multi-layered normal distribution transform for fast and long range scan matching," *Journal of Intelligent & Robotic Systems*, pp. 1–24, 2013.

[16] H. Hong and B. Lee, "Key-layered normal distributions transform for point cloud registration," *Electronics Letters*, vol. 51, no. 24, pp. 1986–1988, 2015.

[17] Q. Li, R. Xiong, and T. Vidal-Calleja, "A gmm based uncertainty model for point clouds registration," *Robotics and Autonomous Systems*, vol. 91, pp. 349–362, 2017.

[18] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, 2012, pp. 3354–3361.

[19] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "Octomap: An efficient probabilistic 3d mapping framework based on octrees," *Autonomous Robots*, vol. 34, no. 3, pp. 189–206, 2013.