FSM: correspondenceless scan-matching of panoramic 2D range scans

Alexandros Filotheou, Georgios D. Sergiadis, Antonis G. Dimitriou

***Abstract*— Recent years have seen the introduction of more affordable but less accurate 2D range sensors whose field of view is 2π. Scan-matching with these has been insufficiently researched, while being a challenge due to these sensors' increased measurement uncertainty. This paper proposes a real-time method for matching scans extracted from panoramic 2D LIDAR sensors. The method leverages properties of the Fourier transform which arise due to the periodicity of the range signal. Matching is performed in a correspondenceless manner. The proposed method outperforms established scan-matching methods in terms of pose accuracy and robustness in tests on public domain data, and over noise levels of commercially available sensors. The source code is available for download.**

***Index Terms*— Scan-matching, localisation, panoramic LIDAR**

I. Introduction

Consider a robot capable of motion, equipped with a Light Detection and Ranging sensor (LIDAR), capturing a measurement $\mathcal{S}_0$ at time t_0 from pose $\boldsymbol{p}_0$ in some reference frame. The robot then moves to pose $\boldsymbol{p}_1$ at time t_1 at which time it captures measurement $\mathcal{S}_1$. Provided overlap between the two scans, estimating the rigid-body transformation $\boldsymbol{T}$ that projects the endpoints of $\mathcal{S}_1$ to those of $\mathcal{S}_0$ with the least error is known as scan-matching. The solution to the scan-matching problem is central to methods of Localisation [1], Navigation [2], and Simultaneous Localisation and Mapping (SLAM) [3], [4], as $\boldsymbol{T}$ is the rigid-body transformation $\boldsymbol{p}_1 - \boldsymbol{p}_0$: i.e. the solution to scan-matching provides localisation information at time t_1, relative to $\boldsymbol{p}_0$. For this reason, along with the high measurement accuracy of LIDAR sensors, scan-matching is also used as a means to improving, providing, or substituting odometric measurements (where available; fig. 1), as the latter are prone to unbounded and unpredictable tire and wheel slippage [5], [6].

LIDAR sensors with a field of view of 360°, i.e. panoramic sensors, were for years constrained to high price ranges, and most provided 3D measurements. Therefore research on scan-matching with 2D LIDAR sensors mostly focused on non-panoramic sensors, with scan matching methods being used without distinction with regard to field of view. In recent years, however, price-appealing 2D LIDAR sensors have emerged, but at the cost of increased measurement uncertainty. The introduction of these sensors warrants targeted research into scan-matching with the use of panoramic LIDAR sensors, due to (a) the afforded periodicity of the range signal, and (b) the need of addressing the high levels of measurement noise with regard to the transformation errors of current scan-matching algorithms.

This work was supported by the European Union and Greek National Funds through the Operational Program Competitiveness, Entrepreneurship, and Innovation, under the call Research Create Innovate under Project T2EDK-02000. Corresponding author: Alexandros Filotheou, alefilot@auth.gr. The authors are with the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

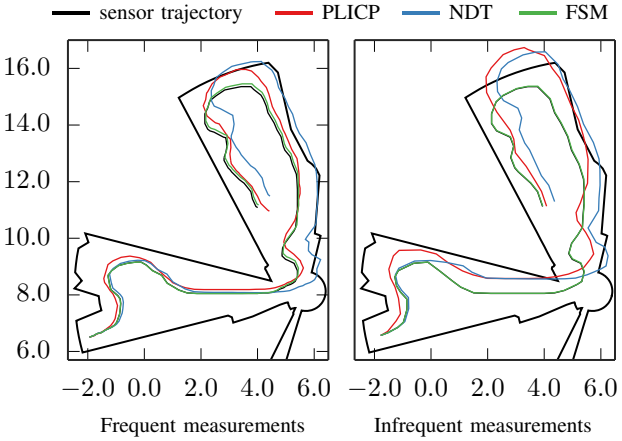

Fig. 1: Scan-matching as "laser odometry": the robot moves from the lower left portion of the environment to the upper right, capturing 2D range scans along its trajectory. The coloured routes show the estimated path of the robot derived from each method. The proposed method's error is invariant to angular and locational displacement

This paper introduces a real-time method specifically targeting the matching of 2D panoramic range scans. Its errors are largely invariant to angular and locational displacement for a given level of measurement noise. The central contributions of the paper are:

- To the best of the author's knowledge, the first method explicitly addressing the matching of panoramic 2D range scans that operates without establishing correspondences between input scans
- The extrication from the need of a prior transformation estimate
- The introduction of a method that aims at reducing the orientation error to lower than the sensor's angle increment compared to relevant prior work
- The parameter set needed by the proposed method is intuitive, smaller than those of established methods, and trades accuracy for execution time
- The thorough evaluation of the proposed method against five established scan-matching algorithms in common use, with benchmark datasets, using measurement noise levels from common-use, commercially available sensors

The rest of the paper is structured as follows. In section II

necessary notions are defined, and the problem of matching panoramic 2D range scans is formulated. A brief review of methods matching 2D range scans is given in section III. Section IV provides an analysis of the proposed method. The experimental setup and are illustrated in section V. Section VI gives characterisations of the proposed method compared to state of the art methods. Section VII concludes this paper.

II. Definitions and Problem Formulation

Definition I. *Definition of a range scan captured from a conventional 2D LIDAR sensor.* A conventional 2D LIDAR sensor provides a finite number of ranges, i.e. distances to objects within its range, on a horizontal cross-section of its environment, at regular angular and temporal intervals, over a defined angular range [7]. A range scan $\mathcal{S}$, consisting of N_s rays over an angular range λ, is an ordered map $\mathcal{S} : \Theta \rightarrow \mathbb{R}_{\geq 0}$, $\Theta = \{\theta_n \in [-\frac{\lambda}{2}, +\frac{\lambda}{2}) : \theta_n = -\frac{\lambda}{2} + \lambda\frac{n}{N_s}, n = 0, 1, \ldots, N_s - 1\}$. Angles θ_n are expressed relative to the sensor's heading, in the sensor's frame of reference. The angular distance between two consecutive rays is the sensor's angle increment $\gamma \triangleq \lambda / N_s$.

Definition II. *Definition of a map-scan.* A map-scan is a virtual scan that encapsulates the same pieces of information as a scan derived from a physical sensor. Only their underlying operating principle is different due to the fact the map-scan refers to distances to the boundaries of a point-set, referred to as the map, rather than within a real environment. A map-scan is derived by means of locating intersections of rays emanating from the estimate of the sensor's pose estimate and the boundaries of the map.

Problem I. Let a mobile robot, capable of motion in the $x-y$ plane, be equipped with a coplanarly mounted range scan sensor emitting N_s rays. Let also the following be available or standing:

- The angular range of the range sensor is 360°
- A 2D range scan $\mathcal{S}_0$, captured at time t_0
- A 2D range scan $\mathcal{S}_1$, captured at $t_1 > t_0$

Then the objective is estimating the 3D rigid-body transformation $\boldsymbol{T} = (\Delta x, \Delta y, \Delta\theta)$ which, when applied to the endpoints of $\mathcal{S}_1$, aligns them to those of $\mathcal{S}_0$ with the least error. Equivalently, roto-translation $\boldsymbol{T}$ corresponds to the relative motion of the sensor from the pose where it captured $\mathcal{S}_0$ to the pose from which it captured $\mathcal{S}_1$.

III. Prior Work

Scan-matching with the use of a 2D LIDAR sensor began with Iterative Dual Correspondences (IDC) [8], an algorithm incorporating elements of the Iterative Closest Point (ICP) algorithm [9]. The latter and its variants, e.g. [10]–[13], have become the de facto scan-matching algorithms in 2D and 3D settings, with research using ICP being still ongoing [14]–[17]. In particular, PLICP [12] has been widely adopted due to its increased accuracy among ICP variants, and the availability of its source code. ICP and its variants, however, exhibit varying performance [18], limited by the noise level in the input scans, the choice of prior, and the configuration of the parameters governing their response. The vast majority of all matching methods adopt ICP's approach of establishing correspondences between the two input scans, using various assumptions, mechanisms (e.g. points to distributions [19], [20]), and types of sources (e.g. features instead of points; a detailed review of scan-matching methods may be found in [21]). The major problem with establishing correspondences is that the process becomes more inefficient and error-prone as measurement noise increases.

By contrast, the method introduced in this paper does not operate by establishing correspondences. Its rotational component is most akin to those of [22] and [23]. They use Phase-Only Matched Filtering (POMF) [24] in both rotation and translation components; the former in one dimension and the latter in two dimensions. In the latter, the requirements for a real-time solution and adequate accuracy cannot be fulfilled simultaneously due to the inability to balance high grid resolution (and therefore high accuracy) with regular sensor updates. The former alleviates this limitation by operating in one dimension, but suffers from the same causes, namely, discretisation errors. Whereas the latter is dependent on the grid's resolution, the former is dependent on the sensor's immutable angle increment. In both methods, both the rotational and translational components are affected, but no mitigation technique is employed to decrease the errors of either their components. By contrast, the method introduced in this paper addresses all the above issues by (a) aiming to extricate the orientation error from the sensor's angle increment, (b) employing a continuous-space translation method, and (c) fulfilling the real-timeness constraint.

IV. Approach

Problem I is iteratively decomposed into two disjunctive sub-problems. The first is estimating the relative orientation of $\mathcal{S}_1$ with respect to $\mathcal{S}_0$ under the assumption that both are captured from the same location (subsection IV-A). In order to do so, $\mathcal{S}_0$ is first transformed into a point-set (the map) by projection to the 2D plane. A number of map-scans are then captured from the projection origin, at consecutive angular intervals. The intervals' sizes are smaller than the sensor's angle increment in order for the magnitude of the angular approximation error to be lower than the angle increment. The map-scans are then compared against $\mathcal{S}_1$ for similarity in the POMF matching sense. The output is an orientation difference estimate between each map-scan and $\mathcal{S}_1$. In order to increase discernibility between the least erroneous orientation estimate and the other estimates, the pose estimate is updated with each orientation estimate, and given over to the translation correction system, whose resulting location error is proportional to the poses's orientation error, for one iteration. The similarity between map-scans captured from these estimates within the map and $\mathcal{S}_1$ is captured in a per-ray range error metric. The final orientation estimate of $\mathcal{S}_1$ is that of the pose estimate whose map-scan scores the lowest error according to this metric.

The second sub-problem is estimating the relative translational displacement of $\mathcal{S}_1$ with respect to $\mathcal{S}_0$ under the assumption that both are captured from poses of the same orientation (subsection IV-B). The solution is given via scan-to-map-scan matching by transforming $\mathcal{S}_0$ to a point-set and matching map-scans, derived from the pose estimate of $\mathcal{S}_1$ within the point-set, to $\mathcal{S}_1$ itself. The matching between the two is derived from first principles given the scans' periodicity and homorientedness; it is facilitated by iteratively updating the position estimate by a vector which is a function of the first term of the DFT of the difference between $\mathcal{S}_1$ and map-scans derived from the point-set of $\mathcal{S}_0$ [25].

In order to solve the joint problem of estimation of position and orientation, these two methods are then joined in tandem in an iterative process (subsection IV-C). At each step the angular intervals at which map-scans are captured in the orientation correction process are progressively shortened. This facilitates finer angular and positional approximation of the sensor's true pose.

A. Estimation of Relative Orientation

Let the assumptions of Problem I be standing. Assume that the two scans were captured from the same location but from different orientations. Denoting with $\mathcal{F}\{\mathcal{S}\}$ the Discrete Fourier Transform (DFT) of signal $\mathcal{S}$, with $\mathcal{F}^{-1}\{\mathcal{S}\}$ its inverse, with $\boldsymbol{c}^*$ the conjugate of complex $\boldsymbol{c}$, and with $|\boldsymbol{c}|$ its magnitude, calculate $Q_{\mathcal{S}_0,\mathcal{S}_1}$:

$$Q_{\mathcal{S}_0,\mathcal{S}_1} \triangleq \frac{\mathcal{F}\{\mathcal{S}_0\}^* \cdot \mathcal{F}\{\mathcal{S}_1\}}{|\mathcal{F}\{\mathcal{S}_0\}| \cdot |\mathcal{F}\{\mathcal{S}_1\}|} \tag{1}$$

on the basis that if space is sampled sufficiently densely, for $k, \xi \in \mathbb{Z}$: $k, \xi \in [0, N_s - 1]$:

$$\mathcal{S}_0[k] \simeq \mathcal{S}_1[(k-\xi) \mod N_s] \Leftrightarrow$$
$$\mathcal{F}\{\mathcal{S}_0\}(u) \simeq e^{-j2\pi\xi u/N_s} \cdot \mathcal{F}\{\mathcal{S}_1\}(u)$$

and, therefore, since $2\pi\dfrac{\xi}{N_s} = \xi\gamma$: $Q_{\mathcal{S}_0,\mathcal{S}_1}(u) \simeq e^{-j\xi\gamma u}$.

The inverse of $Q_{\mathcal{S}_0,\mathcal{S}_1}$ is a Kronecker δ-function $q_{\mathcal{S}_0,\mathcal{S}_1} = \mathcal{F}^{-1}\{Q_{\mathcal{S}_0,\mathcal{S}_1}\}$ centered at $\xi = \arg\max_u q_{\mathcal{S}_0,\mathcal{S}_1}(u)$. If the difference in orientation between the two scans is $\Delta\theta$, then $\Delta\theta = \xi\gamma + \delta\theta$, where $\mod(|\delta\theta|, \gamma) = \lambda \in [0, \frac{\gamma}{2}]$. Therefore for a given number of emitted rays N_s there remains an unresolved orientation error $|\delta\theta| \leq \gamma/2$. The contribution of this error to the scan-matching error is two-fold, as its existence is also propagated to the location estimation method. A method for further reduction of the orientation error is presented in the following.

Let $\mathcal{S}_0$ be projected onto the $x-y$ plane around an arbitrary but fixed pose $\boldsymbol{s}(x_s, y_s, \theta_s)$, producing point-set $\boldsymbol{M}_R$. $\boldsymbol{M}_R$ will hereafter be referred to as the map. Then compute 2^ν map-scans (def. II) $\mathcal{S}_0^k$, $k = 0, \ldots, 2^\nu - 1$, starting from orientation θ_s, at $\gamma/2^\nu$ angular increments. Then the orientation estimation process is carried out once between $\mathcal{S}_1$ and map scan $\mathcal{S}_0^k$ taken from orientation $\theta_0^k = \theta_s + k \cdot \gamma/2^\nu$, for a total of 2^ν times. An alignment metric between the k-th map scan $\mathcal{S}_0^k$ and scan $\mathcal{S}_1$ is computed according to

$$\mathrm{PD}_k = \frac{2\max q_{\mathcal{S}_0^k,\mathcal{S}_1}}{\max q_{\mathcal{S}_0^k,\mathcal{S}_0^k} + \max q_{\mathcal{S}_1,\mathcal{S}_1}} \tag{2}$$

The Percent Discrimination metric $\mathrm{PD}_k \in [0, 1]$, and is proportional to the degree of alignment between map-scan $\mathcal{S}_0^k$ and scan $\mathcal{S}_1$, across all 2^ν map-scans $\mathcal{S}_0^k$. The above analysis is the equivalent of the 2D Fourier-Mellin Invariant matching in one dimension [24].

Let now K denote the index of the k-th map scan $\mathcal{S}_0^K$ scoring the highest PD_k: $\mathrm{PD}_K = \max\{\mathrm{PD}_k\}$, $k = 0, \ldots, 2^\nu - 1$. Let also Ξ denote the integer multiple of angle increments γ by which $\mathcal{S}_0^K$ should be rotated counter-clockwise in order to achieve PD_K: $\Xi = \arg\max_u q_{\mathcal{S}_0^K,\mathcal{S}_1}$. Then the sensor's orientation difference becomes $\Delta\theta = \Xi\gamma + K \cdot \gamma/2^\nu + \delta\theta'$.

If map-scans $\mathcal{S}_0^k$ were computed by raycasting the map of the environment instead of $\boldsymbol{M}_R$ then the residual and unresolved orientation error $|\delta\theta'| \in [0, \gamma/2^{1+\nu}]$. In this case, however, $\boldsymbol{M}_R$ is an approximation of the environment's map in the locality of $\boldsymbol{p}_0$. Depending on the magnitude of the sensor's angle increment and the arbitrariness of the environment, this approximation may be viewed as induced local perturbations in the map of the environment. This holds true in the general case as well, where $\mathcal{S}_0$ and $\mathcal{S}_1$ are captured from different locations. Therefore attaining $|\delta\theta'| \leq \gamma/2^{1+\nu}$ may not always be possible for all combinations of environments and sensor angle increments.

B. Estimation of Relative Location

Let the assumptions of Problem I hold. Assume now that $\mathcal{S}_0$ and $\mathcal{S}_1$ were captured from different positions in the same environment but with the same orientation relative to a fixed reference frame. Let $\mathcal{S}_0$ be projected onto the $x-y$ plane around pose $\boldsymbol{s}(0,0,0)$, producing point-set $\boldsymbol{M}_L$. Assuming that $\mathcal{S}_1$ was captured in a neighbourhood of $\mathcal{S}_0$, then $\boldsymbol{M}_L$ is a perturbed local map of the environment with respect to sensor measurement $\mathcal{S}_1$. Aside from measurement noise, this perturbation manifests due to the finiteness of the sensor's angle increment and to the fact that different portions of the environment are perceptible and therefore measurable from different locations [5]. The nature of these perturbations on map-scans captured within $\boldsymbol{M}_L$ is additive and finite. Under these assumptions the problem of (scan-)matching scan $\mathcal{S}_1$ to scan $\mathcal{S}_0$ may be transformed into a problem of scan–to–map-scan matching, where the aim is registering scan $\mathcal{S}_1$ to map $\boldsymbol{M}_L$: i.e. estimating the pose $\boldsymbol{p}_1$ from where $\mathcal{S}_1$ was captured within $\boldsymbol{M}_L$. Theorem I [25] guarantees that the error of the location estimate between the poses from which the two scans were captured is bounded in a neighbourhood of the origin, when $\hat{\boldsymbol{p}}_1 = \boldsymbol{s}$.

Theorem I. *Let a panoramic 2D range scan $\mathcal{S}_1$ be captured from a physical range sensor from unknown pose $\boldsymbol{p}_1 = (\boldsymbol{l}_1, \theta_1)$, $\boldsymbol{l}_1 = (x_1, y_1)$. Let $\boldsymbol{M}_L$ be the map, i.e. the projection of panoramic 2D range scan $\mathcal{S}_0$ from $\boldsymbol{s}(0,0,0)$. Let a pose estimate $\hat{\boldsymbol{p}}_1 = (\hat{\boldsymbol{l}}_1, \hat{\theta}_1)$ reside in the neighbourhood of $\boldsymbol{p}_1$ in the map's frame of reference. Additionally, let $\hat{\theta}_1 = \theta_1$.*

Assume that $\mathcal{S}_0$ and/or $\mathcal{S}_1$ are affected by additive, bounded disturbances. Then, treating the estimate of the location of the sensor as a state variable $\hat{\boldsymbol{l}}_1[k] = [\hat{x}_1[k], \hat{y}_1[k]]^\top$ and updating it according to the difference equation $\hat{\boldsymbol{l}}_1[k+1] = \hat{\boldsymbol{l}}_1[k] + \boldsymbol{u}[k]$, where $\hat{\boldsymbol{l}}_1[0] = \hat{\boldsymbol{l}}_1 = [\hat{x}_1, \hat{y}_1]^\top$, i.e. the supplied initial location estimate, and $\boldsymbol{u}$ is

$$\boldsymbol{u}[k] = \frac{1}{N_s} \begin{bmatrix} \cos\hat{\theta}_1 & \sin\hat{\theta}_1 \\ \sin\hat{\theta}_1 & -\cos\hat{\theta}_1 \end{bmatrix} \begin{bmatrix} X_{1,r}\big(\mathcal{S}_1, \mathcal{S}_0|_{\hat{\boldsymbol{p}}_1[k]}\big) \\ X_{1,i}\big(\mathcal{S}_1, \mathcal{S}_0|_{\hat{\boldsymbol{p}}_1[k]}\big) \end{bmatrix} \quad (3)$$

where $X_{1,r}(\cdot)$ and $X_{1,i}(\cdot)$ are, respectively, the real and imaginary parts of the complex quantity X_1:

$$X_1\big(\mathcal{S}_1, \mathcal{S}_0|_{\hat{\boldsymbol{p}}_1[k]}\big) = \sum_{n=0}^{N_s-1} \left(\mathcal{S}_1[n] - \mathcal{S}_0[n]|_{\hat{\boldsymbol{p}}_1[k]}\right) \cdot e^{-i\frac{2\pi n}{N_s}} \quad (4)$$

where $\mathcal{S}_1[n]$ and $\mathcal{S}_0[n]|_{\hat{\boldsymbol{p}}[k]}$ are, respectively, the ranges of the n-th ray of scan $\mathcal{S}_1$, and map-scan $\mathcal{S}_0|_{\hat{\boldsymbol{p}}[k]}$ captured via raycasting the map $\boldsymbol{M}_L$ from $\hat{\boldsymbol{p}}_1[k] = (\hat{\boldsymbol{l}}_1[k], \hat{\theta}_1)$— then $\hat{\boldsymbol{l}}_1[k]$ is uniformly bounded for $k \geq k_0$ and uniformly ultimately bounded in a neighbourhood of $\boldsymbol{l}_1$. Its size depends on the suprema of the disturbance corrupting the range measurements of the two scans.

Remark I. Without loss of generality, subsequent to the application of Theorem I, the location error is proportional to the orientation error.

Let $\hat{\boldsymbol{p}}'_1$ denote the resulting pose estimate of $\boldsymbol{p}_1$ in $\boldsymbol{M}_L$. Then $\hat{\boldsymbol{T}} = \hat{\boldsymbol{p}}'_1 - \boldsymbol{s} = \hat{\boldsymbol{p}}'_1$ is the estimate of the 3D rigid transformation of the sensor as it moved from the pose where it captured $\mathcal{S}_0$ to that where it captured $\mathcal{S}_1$.

C. Joint Estimation of Relative Orientation and Location

The previous two sections describe two methods of how it is possible to (a) estimate the relative orientation between two panoramic 2D range scans when both are captured from the same position but from different orientations, and (b) estimate their relative location when both are captured from the same sensor orientation but from different locations. In the general case, however, no equality stands. The following analysis describes how these two methods are combined in tandem in order to solve Problem I.

Let the assumptions of Problem I hold. Then denote by $\boldsymbol{M}$ the point-set that is the result of the projection of range scan $\mathcal{S}_0$ to the $x-y$ plane around $\boldsymbol{s}(0,0,0)$. Then the objective is estimating the pose $\boldsymbol{p}_1$ from where $\mathcal{S}_1$ was captured relative to $\boldsymbol{s}$ by way of registering $\mathcal{S}_1$ to map $\boldsymbol{M}$.

Given an input pose estimate $\hat{\boldsymbol{p}}_1(\hat{x}_1, \hat{y}_1, \hat{\theta}_1)$, range scan $\mathcal{S}_1$, the map $\boldsymbol{M}$, and a sampling degree ν, the One-step Pose Estimation system (fig. 2) first calculates 2^ν pose estimates of $\boldsymbol{p}_1$: $\boldsymbol{P}_{OC} = \{(\hat{x}_1, \hat{y}_1, \hat{\theta}_1^k)\}$, $k = 0, \ldots, 2^\nu - 1$, according to the orientation estimation method described in section IV-A. The initial pose estimate of $\boldsymbol{p}_1$ is $\hat{\boldsymbol{p}}_1 = \boldsymbol{s}$. If scans $\mathcal{S}_0$ and $\mathcal{S}_1$ were captured from the same location, then the Percent Discrimination metric (eq. 2) would suffice in serving as an accurate determinant of the orientation of $\boldsymbol{p}_1$. In practice, however, the ranking provided by the Percent Discrimination

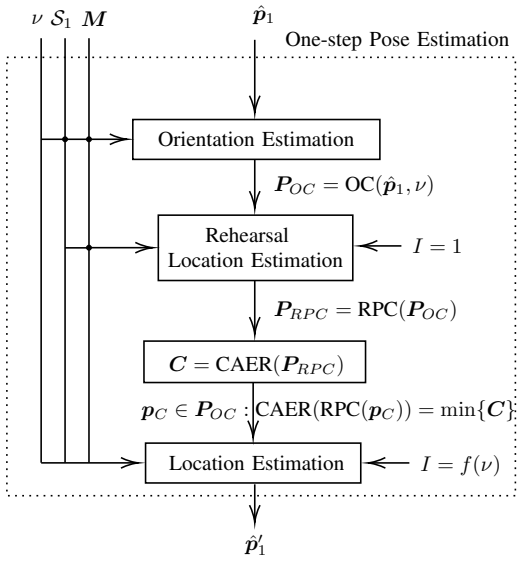

Fig. 2: FSM iteratively invokes the One-step Pose Estimation method. Given a pose estimate of where scan $\mathcal{S}_1$ was captured within $\boldsymbol{M}$, the method attempts to register $\mathcal{S}_1$ to $\boldsymbol{M}$ by estimating first its relative orientation and then its location with respect to the input pose estimate

metric is confounded by the incoincidence of the two locations. In order to mitigate this effect, each pose estimate in $\boldsymbol{P}_{OC}$ is given over to the Position Estimation system, where the position of each pose estimate is displaced once ($I = 1$), according to the method described in section IV-B. This operation produces the pose set $\boldsymbol{P}_{RPC} = \{(\hat{x}_1^k, \hat{y}_1^k, \hat{\theta}_1^k)\}$, $|\boldsymbol{P}_{RPC}| = 2^\nu$. The purpose of this operation is for it to provide an advance view of the next step of location estimation: the less rotationally misaligned a pose estimate of $\boldsymbol{p}_1$ is, the less it will diverge in terms of orientation and hence position with respect to $\boldsymbol{p}_1$ once inputted to the position estimation system (remark I). This divergence is captured by the Cumulative Absolute Error per Ray (CAER) metric:

$$\text{CAER}_k = \sum_{n=0}^{N_s-1} \left| \mathcal{S}_1[n] - \mathcal{S}_0^k[n] \Big|_{(\hat{x}_1^k, \hat{y}_1^k, \hat{\theta}_1^k)} \right| \quad (5)$$

where $\mathcal{S}_0^k$ is the map-scan captured from $(\hat{x}_1^k, \hat{y}_1^k, \hat{\theta}_1^k)$, $k = 0, \ldots, 2^\nu - 1$, within $\boldsymbol{M}$. The CAER metric (fig. 3) encodes at the same time a degree of alignment of position and orientation between its two input scans. By rehearsing the position estimation of each pose estimate in $\boldsymbol{P}_{OC}$ and capturing the CAER for each of its displaced pose estimates in $\boldsymbol{P}_{RPC}$, it is possible to establish a pose error rank between pose estimates in $\boldsymbol{P}_{OC}$ and simultaneously retain only one pose estimate for the next iteration of the One-step Pose Estimation method.The pose estimate $\boldsymbol{p}_C \in \boldsymbol{P}_{OC}$ which, when translated once, records the minimum CAER among all similarly-treated pose estimates in $\boldsymbol{P}_{OC}$ is inputted to the Position Estimation method proper. The number of translation iterations I it undergoes is an increasing function in the degree of map sampling ν. The Position Estimation

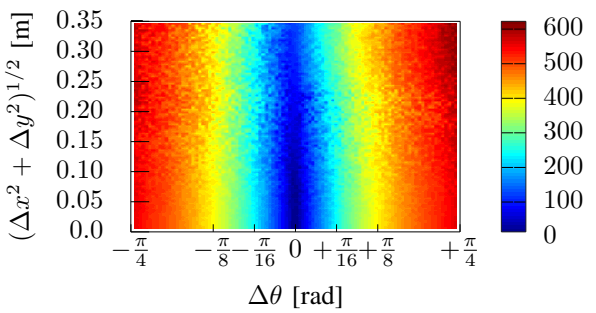

Fig. 3: A profile of the CAER metric (eq. (5)) from 10^6 pairs of sample scans, depending on the distance $(\Delta x^2 + \Delta y^2)^{1/2}$ and relative orientation $\Delta\theta$ of the poses from where the two scans were captured. Pose estimates closer to the true pose in terms of orientation (a) exhibit lower CAER values than those further away from it and (b) produce lower location errors once inputted to the Location Estimation system

system produces $\hat{\boldsymbol{p}}'_1$, which is then fed back to the Orientation Estimation system in the form of a new pose estimate of $\boldsymbol{p}_1$: $\hat{\boldsymbol{p}}_1 \leftarrow \hat{\boldsymbol{p}}'_1$. In practice, the pose set $\boldsymbol{P}_{OC}$ is supplemented with one pose whose location component is equal to $\hat{\boldsymbol{p}}_1$ and whose orientation is equal to the orientation of $\boldsymbol{p}_C$ that produces the minimum CAER over time. This addition introduces a form of memory to the system, which assists it in avoiding divergence and which, therefore, benefits speed of execution.

Given pose $\hat{\boldsymbol{p}}_1$, range scan $\mathcal{S}_1$, and the map $\boldsymbol{M}$, the pose estimation method proposed iteratively invokes the One-step Pose Estimation process until a set of termination conditions is met. Denoting the former by FSM (Fourier Scan Matching), FSM starts off with an initial degree of sampling the map $\nu = \nu_{\min}$. The input pose estimate $\hat{\boldsymbol{p}}_1$ is processed by the One-step Pose Estimation process, and its output $\hat{\boldsymbol{p}}'_1$ is examined with regard to Recovery and Convergence conditions. If the resulting pose estimate falls outside of the map $\boldsymbol{M}$ then a new pose estimate is generated from the initially supplied pose estimate $\boldsymbol{s}$, and the process is reset. If no significant pose estimate correction is observed $\|\hat{\boldsymbol{p}}'_1 - \hat{\boldsymbol{p}}_1\|_2 < \varepsilon_{\delta p}$, then the degree of map sampling ν is increased. Its increase serves as a means of reducing the orientation and hence the position estimate error further. Otherwise, the One-step Pose Estimation process is iterated until a maximum degree of map sampling is reached $\nu = \nu_{\max}$, at which point FSM terminates. Its output is $\hat{\boldsymbol{p}}'_1$, which is the pose estimate of $\boldsymbol{p}_1$ in the frame of reference of $\boldsymbol{M}$. The roto-translation $\hat{\boldsymbol{T}} = \hat{\boldsymbol{p}}'_1 - \boldsymbol{s} = \hat{\boldsymbol{p}}'_1$ is the estimate of the sensor's true motion $\boldsymbol{T}$. Algorithm I describes FSM in pseudocode.

Algorithm I FSM

Input: $\mathcal{S}_0$, $\mathcal{S}_1$, γ, $[\nu_{\min}, \nu_{\max}, I, \varepsilon_{\delta p}]$
Output: $\hat{\boldsymbol{p}}_1(\hat{x}_1, \hat{y}_1, \hat{\theta}_1)$

$\hat{\boldsymbol{p}}_1 \leftarrow (0,0,0)$, $\nu \leftarrow \nu_{\min}$
$\boldsymbol{M} \leftarrow$ `project_to_2d`$(\mathcal{S}_0, \hat{\boldsymbol{p}}_1)$
while $\nu \leq \nu_{\max}$ **do**
  $\boldsymbol{P}_{RPC}, \boldsymbol{C} \leftarrow \{\varnothing\}$
  for $k \leftarrow 0, \ldots, 2^\nu - 1$ **do**
    $\mathcal{S}_0^k \leftarrow$ `scan_map`$(\boldsymbol{M}, \hat{\boldsymbol{p}}_1 + [0, 0, k\gamma 2^{-\nu}])$
    $\xi_1^k \leftarrow \arg\max \mathcal{F}^{-1}\{Q_{\mathcal{S}_0^k, \mathcal{S}_1}\}$ (eq. 1)
    $\hat{\boldsymbol{p}}_1^k \leftarrow \hat{\boldsymbol{p}}_1 + [0, 0, \xi_1^k \gamma + k\gamma 2^{-\nu}]$ (sec. IV-A)
    $\hat{\boldsymbol{p}}_1^k \leftarrow$ `translate`$(\hat{\boldsymbol{p}}_1^k, \boldsymbol{M}, \mathcal{S}_1, 1)$ (sec. IV-C)
    $\boldsymbol{P}_{RPC} \leftarrow \{\boldsymbol{P}_{RPC}, \hat{\boldsymbol{p}}_1^k\}$
    $\boldsymbol{C} \leftarrow \{\boldsymbol{C},$ `CAER`$(\mathcal{S}_1,$ `scan_map`$(\boldsymbol{M}, \hat{\boldsymbol{p}}_1^k))\}$ (eq. 5)
  end for
  $\hat{\boldsymbol{p}}'_1 \leftarrow \boldsymbol{P}_{RPC}[\arg\min \boldsymbol{C}]$
  $\hat{\boldsymbol{p}}'_1 \leftarrow$ `translate`$(\hat{\boldsymbol{p}}'_1, \boldsymbol{M}, \mathcal{S}_1, I)$; (sec. IV-B)
  if $\|\hat{\boldsymbol{p}}'_1 - \hat{\boldsymbol{p}}_1\| < \varepsilon_{\delta p}$ **then**
    $\nu = \nu + 1$
  end if
  if $\hat{\boldsymbol{p}}'_1$ not in $\boldsymbol{M}$ **then**
    generate new $\hat{\boldsymbol{p}}_1$; $\nu \leftarrow \nu_{\min}$; continue
  end if
  $\hat{\boldsymbol{p}}_1 \leftarrow \hat{\boldsymbol{p}}'_1$
end while
return $\hat{\boldsymbol{p}}_1$

V. Results

A. Experimental Design

The experimental procedure was conducted using a collection $D = \{D_k\}, k = 1, \ldots, 5$ of five heterogeneous benchmark datasets and sensor properties, courtesy of the Department of Computer Science, University of Freiburg, comprising a total of $|D| = 45402$ scan measurements [26].

For purposes of comparison against state-of-the-art scan-matching methods the experimental procedure is extended to PLICP [12], the Normal Distributions Transform (NDT) scan-matching method [19], and FastGICP [13]. PLICP, NDT, and FastGICP belong to the *established* state-of-the-art methods of scan-matching [17], [27]–[31]. In addition, the experimental procedure is extended to FastVGICP [31] and NDT-PSO [20].

The experimental setup is the following. The rays of each dataset instance $D^d, d = 1, 2, \ldots, |D|$ are first projected to the $x - y$ plane around $\boldsymbol{r}^d$. The dataset's scans are not panoramic, therefore the remaining space is filled with a semicircular arc that joins the scan's two extreme ends. Alternative fashions for closing-off the environment have been found equivalent with respect to the performance of the tested methods. The resulting point-set is regarded as the environment $\boldsymbol{W}^d$ in which the range sensor operates. Then the pose $\boldsymbol{p}_0^d$ from which $\mathcal{S}_0^d$ is captured is generated randomly within the polygon formed by $\boldsymbol{W}^d$. The pose $\boldsymbol{p}_1^d$ from which the sensor captured $\mathcal{S}_1$ is then obtained by perturbing the components of $\boldsymbol{p}_0^d$ with quantities extracted from uniformly distributed error distributions $U_{xy}(-\overline{\delta}_{xy}, \overline{\delta}_{xy})$, $U_\theta(-\overline{\delta}_\theta, \overline{\delta}_\theta)$; $\overline{\delta}_{xy}$, $\overline{\delta}_\theta \in \mathbb{R}_{\geq 0}$.

Range scans $\mathcal{S}_0^d$ and $\mathcal{S}_1^d$ are then computed by locating the intersection points between N_s rays emanating from $\boldsymbol{p}_0^d$ and $\boldsymbol{p}_1^d$, respectively, and the polygon formed by $\boldsymbol{W}^d$ across an angular field of view $\lambda = 2\pi$. The inputs to all algorithms are then set to $\mathcal{S}_0^d$ and $\mathcal{S}_1^d$. Their output is $\boldsymbol{p}_1'^d$. The roto-translation $\hat{\boldsymbol{T}}^d = \boldsymbol{p}_1'^d$ is the estimate of the motion $\boldsymbol{T}^d = \boldsymbol{p}_1^d - \boldsymbol{p}_0^d$ of the range sensor.

For every pose estimate $\boldsymbol{p}_1'^d$ outputted by each algorithm, its offset from the actual pose $\boldsymbol{p}_1^d$ is recorded in the form of the orientation error and the 2-norm position error. In order to test for the performance of the proposed method with use of real sensors, five levels of noise acting on the range measurements of the scans are tested. The range measurements are perturbed by zero-mean normally-distributed noise with standard deviation $\sigma_R \in \{0.01, 0.03, 0.05, 0.10, 0.20\}$ m. The values of tested standard deviations were calculated from commercially available panoramic LIDAR scanners by identifying the magnitude of their reported maximum range errors and dividing it by a factor of three. The rationale is that 99.73% of errors are located within 3σ around the actual range between a ray and an obstacle, assuming errors are distributed normally. The minimum standard deviation is reported for VELODYNE sensors [32]; the rest are reported for price-appealing but disturbance-laden RPLIDAR [33] and YDLIDAR [34] sensors. The size of the input scans was set to $N_s = 360$ rays. The minimum and maximum map oversampling rates of FSM were set to $(2^{\nu_{\min}}, 2^{\nu_{\max}}) = (2^0, 2^3)$. The number of iterations of the translational component at each map sampling degree ν was set at $I = 5\nu$. The orientation convergence threshold was set to $\varepsilon_{\delta p} =$ 10e-5. Maximal displacements $\overline{\delta}_{xy}$ and $\overline{\delta}_\theta$ were chosen as such by prior art tests [12]. For each experiment all algorithms ran for $E = 10$ times across all instances of D; therefore each method underwent a total of $10 \times 45402 \times 6 \times 5 \sim O(10^8)$ experiments. The orientation and position error distributions reported below are those across all $E \cdot |D|$ experiments of the same configuration. All experiments and algorithms were run in C++, on a single thread, on a machine with a CPU frequency of 4.0 GHz. The implementations of PLICP, NDT, FastGICP, FastVGICP, and NDT-PSO were taken from [35].

B. Performance

Figure 4 shows the distribution of rotation and translation errors across all experiments for all tested algorithms. FSM's position and orientation errors are equal to or lower than the most accurate method for each displacement and sensor noise configuration. As displacements and sensor noise levels increase, its errors increase at a lower rate than any tested method. The magnitude of FSM's errors is largely independent of the displacement of the two input scans for a given level of sensor noise (fig. 1). In terms of orientation, 72%-74% of FSM's errors resulted below $\gamma/2^{\nu_{max}+1} = 0.0625$ deg when $\sigma_R = 0.01$ m, and 33%-36% when $\sigma_R = 0.20$ m. The juxtaposition of the six methods' errors at high levels of sensor noise highlight the robustness afforded to FSM by the Discrete Fourier transform and its properties. In terms of execution time, PLICP ranged between 4.8-17.5 ms, NDT 8.1-19.9 ms, FastGICP 3-9 ms, FastVGICP 3.8-6.8 ms, NDT-PSO 190-200 ms, and FSM between 13.2 and 16.7 ms. The measurement frequency of modern LIDAR sensors ranges from 12-20 Hz; therefore FSM runs in real time in modern processors.

VI. Characterisation

Figure 5 depicts a comparison of FSM's alignment process against that of an ICP variant, namely FastGICP, for measurement noise with $\sigma_R = 0.03$ m. The top figure shows the initial configuration between two scans captured from poses differing by $\boldsymbol{T} = (0.35\text{m}, 0.1\text{m}, -60\text{deg})$. While FastGICP progressively reduces the orientation error, FSM nearly eliminates it in one step. From there it provides finer approximations of the true orientation by increasing its angular sampling degree, which in turn allow it to eliminate the positional errors further. The CAER metric succeeds in navigating the error space better than FastGICP's internal error metric, especially at such high orientational and positional displacements. This fact makes FSM suitable in conditions of higher robot velocities (here up to 4.2 m/s and 720 deg/s). FSM converges with an error of $(0.0098\text{m}, 0.08\text{deg})$ in half as many iterations as FastGICP, whose final error is $(0.61\text{m}, 9.94\text{deg})$.

Figure 6 shows the behaviour of FSM and an ICP variant, namely, PLICP, in tests where the sensor's maximum range is limited. Specifically, it shows their mean orientational and positional errors in ten tests, for progressively smaller maximum sensor range, within the two environments depicted in the top row with white colour. The sensor is affected by noise of $\sigma_R = 0.05$ m. The sensor's first position is marked with a blue dot. The maximum displacement between sensor poses is set to $(\overline{\delta}_{xy}, \overline{\delta}_\theta) = (0.05\text{m}, 10\text{deg})$ and $(\overline{\delta}_{xy}, \overline{\delta}_\theta) = (0.20\text{m}, 45\text{deg})$. Altough one would expect that matching without the use of correspondences would fare worse than matching with it, FSM exhibits more robust and more accurate orientational errors than PLICP. In terms of position, FSM's errors are equal to or lower than those of PLICP.

VII. Conclusions and Future Steps

This paper has presented a scan-matching method for panoramic LIDAR sensors. The approach rests on properties of the DFT, which afford it increased robustness and accuracy compared to established scan-matching approaches in the face of measurement noise exhibited by real-life sensors. Since FSM does not rely on correspondences or features, and may operate under missing range information, it is also suitable for unstructured and outdoor environments.

Future work will focus on extension to 3D LIDAR sensors, where the slices not parallel to the ground may provide vertical motion information according to the sine of each slice's pitch angle, and tests on more diverse environments, such as those with high-frequency components.

The C++ code of the proposed method, along with the implementation of the conducted experiments is available at `https://github.com/li9i/fsm`.

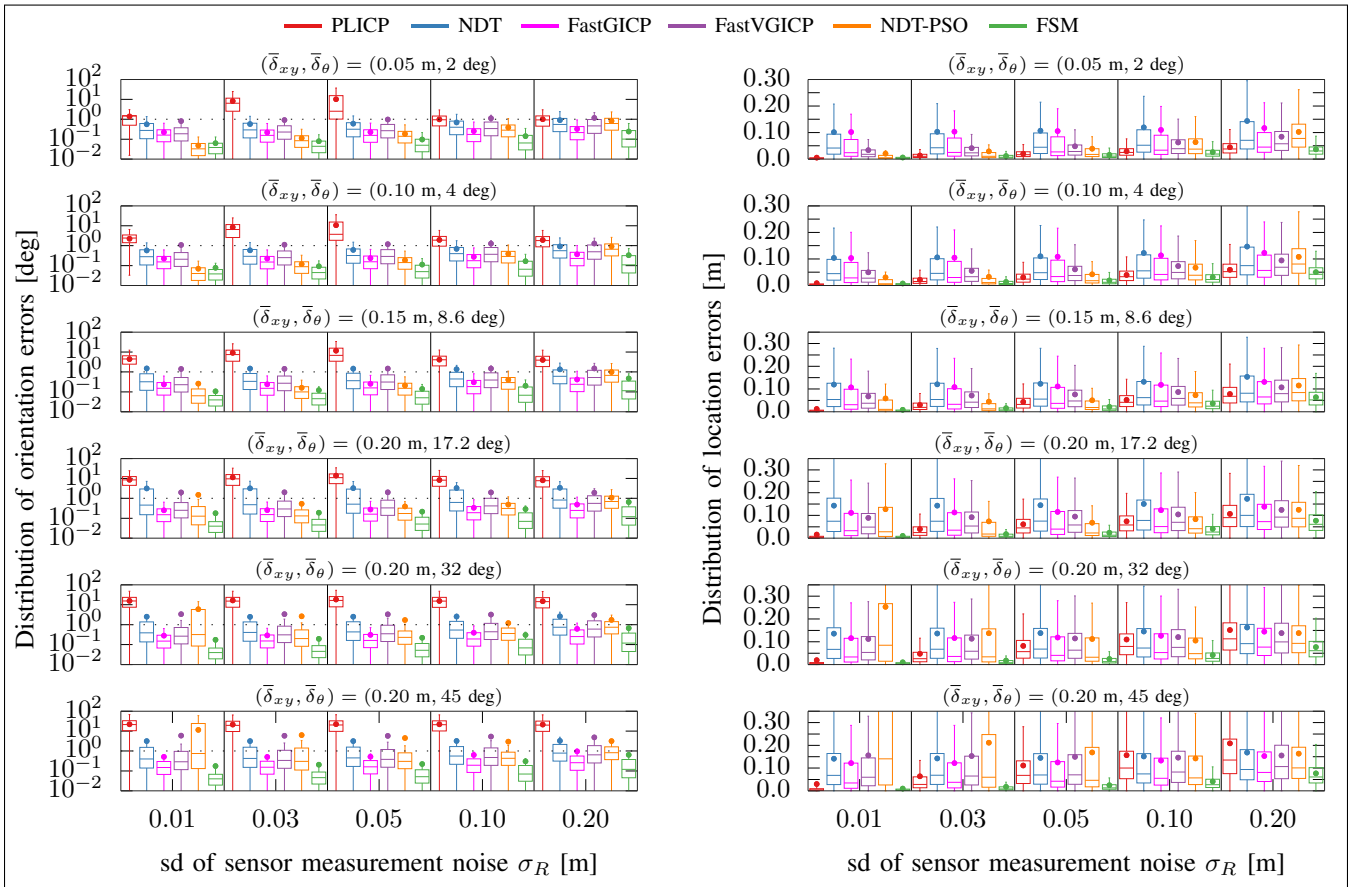

Fig. 4: Distribution of orientation and position errors across a range of maximal positional and orientational displacements, for progressively larger sensor measurement noise levels. Each boxplot represents 10 iterations over $\sum |D_k| \approx 45{\cdot}10^3$ random scan pairs for each configuration, $k = 1, \ldots, 5$. Dots signify mean errors. FSM's errors are largely independent of the initial displacement of scans for a given level of sensor noise

REFERENCES

[1] X. Ju, D. Xu, X. Zhao, W. Yao, and H. Zhao, "Learning Scene Adaptive Covariance Error Model of LiDAR Scan Matching for Fusion Based Localization," in *2019 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, jun 2019, pp. 1789–1796.

[2] N. Kumar and Z. Vamossy, "Laser Scan Matching in Robot Navigation," in *2018 IEEE 12th International Symposium on Applied Computational Intelligence and Informatics (SACI)*. IEEE, may 2018, pp. 000 241–000 246.

[3] H. Zhang, N. Chen, G. Fan, and D. Yang, "An improved scan matching algorithm in SLAM," in *2019 6th International Conference on Systems and Informatics (ICSAI)*. IEEE, nov 2019, pp. 160–164.

[4] E. Pedrosa, A. Pereira, and N. Lau, "Fast Grid SLAM Based on Particle Filter with Scan Matching and Multithreading," in *2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. IEEE, apr 2020, pp. 194–199.

[5] E. Olson, "Real-time correlative scan matching," in *2009 IEEE International Conference on Robotics and Automation*. IEEE, may 2009, pp. 4387–4393.

[6] S. Zhang, L. Xiao, Y. Nie, B. Dai, and C. Hu, "Lidar Odometry and Mapping Based on Two-stage Feature Extraction," in *2020 39th Chinese Control Conference (CCC)*. IEEE, jul 2020, pp. 3966–3971.

[7] M. Cooper, J. Raquet, and R. Patton, "Range Information Characterization of the Hokuyo UST-20LX LIDAR Sensor," *Photonics*, vol. 5, no. 2, p. 12, may 2018.

[8] Feng Lu and Milios, "Robot pose estimation in unknown environments by matching 2D range scans," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94*. IEEE Comput. Soc. Press, 1994, pp. 935–938.

[9] P. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, feb 1992.

[10] S. Pfister, K. Kriechbaum, S. Roumeliotis, and J. Burdick, "Weighted range sensor matching algorithms for mobile robot displacement estimation," in *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, vol. 2. IEEE, pp. 1667–1674.

[11] D. Chetverikov, D. Svirko, D. Stepanov, and P. Krsek, "The Trimmed Iterative Closest Point algorithm," in *Object recognition supported by user interaction for service robots*, vol. 3. IEEE Comput. Soc, pp. 545–548.

[12] A. Censi, "An ICP variant using a point-to-line metric," in *2008 IEEE International Conference on Robotics and Automation*. IEEE, may 2008, pp. 19–25.

[13] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," in *Robotics: Science and Systems V*. Robotics: Science and Systems Foundation, jun 2009.

[14] J. Wang, M. Zhao, and W. Chen, "MIM_SLAM: A Multi-Level ICP Matching Method for Mobile Robot in Large-Scale and Sparse Scenes," *Applied Sciences*, vol. 8, no. 12, p. 2432, nov 2018.

[15] Y. Tian, X. Liu, L. Li, and W. Wang, "Intensity-Assisted ICP for Fast Registration of 2D-LIDAR," *Sensors*, vol. 19, no. 9, p. 2124, may 2019.

[16] Ł. Marchel, C. Specht, and M. Specht, "Testing the Accuracy of the Modified ICP Algorithm with Multimodal Weighting Factors," *Energies*, vol. 13, no. 22, p. 5939, nov 2020.

[17] K. Koide, M. Yokozuka, S. Oishi, and A. Banno, "Voxelized GICP for Fast and Accurate 3D Point Cloud Registration," in *2021 IEEE*

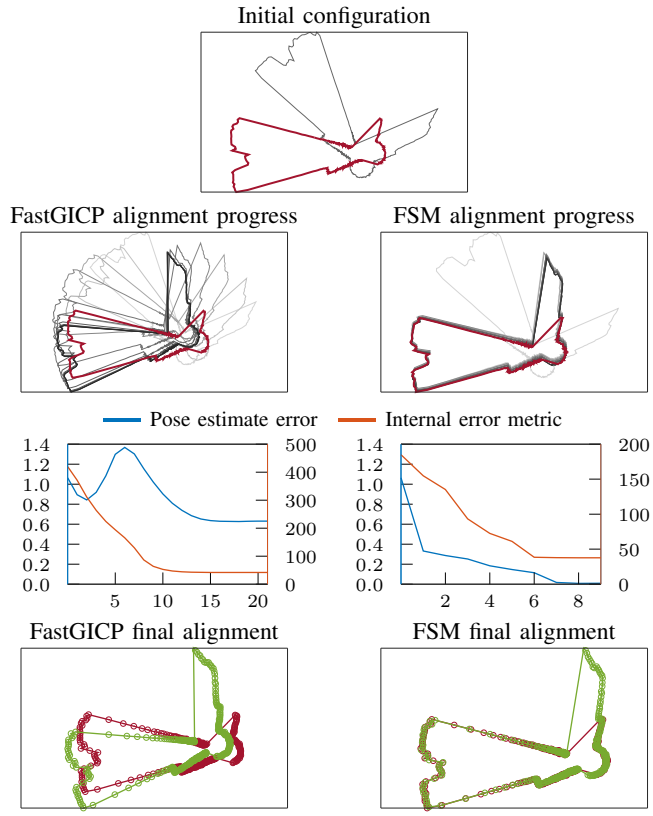

Fig. 5: Comparison of alignment progress of FSM against FastGICP

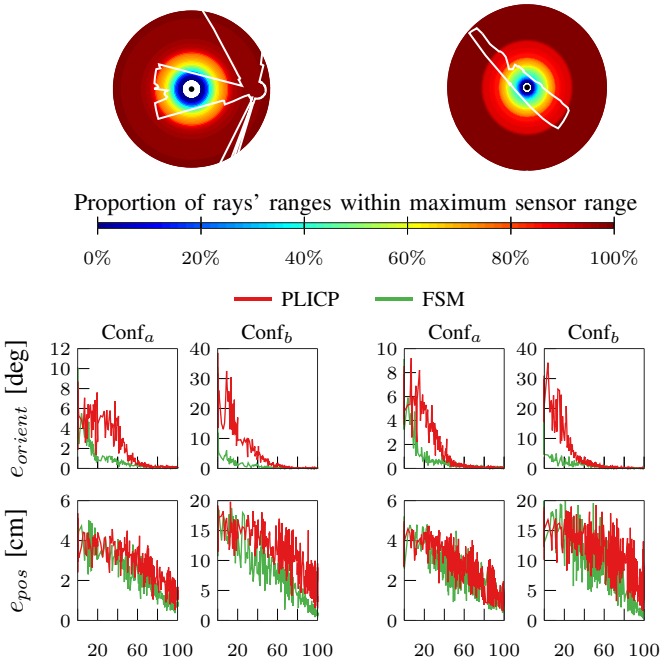

Fig. 6: FSM's performance with respect to receding maximum sensor range, compared to that of PLICP. Conf$_a$ is configuration $(\overline{\delta}_{xy}, \overline{\delta}_{\theta}) = (0.05\text{m}, 10\text{deg})$. Conf$_b$ is $(\overline{\delta}_{xy}, \overline{\delta}_{\theta}) = (0.20\text{m}, 45\text{deg})$

International Conference on Robotics and Automation (ICRA). IEEE, may 2021, pp. 11 054–11 059.

[18] F. Donoso, K. Austin, and P. McAree, "How do ICP variants perform when used for scan matching terrain point clouds?" *Robotics and Autonomous Systems*, vol. 87, pp. 147–161, jan 2017.

[19] P. Biber and W. Strasser, "The normal distributions transform: a new approach to laser scan matching," in *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453)*, vol. 3. IEEE, pp. 2743–2748.

[20] S. Bouraine, A. Bougouffa, and O. Azouaoui, "NDT-PSO, a New NDT based SLAM Approach using Particle Swarm Optimization," in *2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. IEEE, dec 2020, pp. 321–326.

[21] A. Filotheou, E. Tsardoulias, A. Dimitriou, A. Symeonidis, and L. Petrou, "Pose Selection and Feedback Methods in Tandem Combinations of Particle Filters with Scan-Matching for 2D Mobile Robot Localisation," *Journal of Intelligent & Robotic Systems*, vol. 100, no. 3-4, pp. 925–944, dec 2020.

[22] H. Yu, Y. Zeng, and H. Dai, "A Novel Scan Matching Method for Mobile Robot Based on Phase Only Matched Filtering *," in *2018 IEEE International Conference on Information and Automation (ICIA)*. IEEE, aug 2018, pp. 391–394.

[23] G. Jiang, L. Yin, G. Liu, W. Xi, and Y. Ou, "FFT-Based Scan-Matching for SLAM Applications with Low-Cost Laser Range Finders," *Applied Sciences*, vol. 9, no. 1, p. 41, dec 2018.

[24] Qin-Sheng Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 12, pp. 1156–1168, 1994.

[25] A. Filotheou, "Correspondenceless scan-to-map-scan matching of homoriented 2D scans for mobile robot localisation," *Robotics and Autonomous Systems*, vol. 149, mar 2022.

[26] "Datasets used in the experimental procedure." [Online]. Available: http://ais.informatik.uni-freiburg.de/slamevaluation/datasets.php

[27] P. Xu, F. Gu, Z. Song, and J. Li, "Non-iterative multiple data registration method based on the motion screw theory and trackable features," in *2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE, aug 2018, pp. 2428–2432.

[28] H. Sobreira, C. M. Costa, I. Sousa, L. Rocha, J. Lima, P. C. M. A. Farias, P. Costa, and A. P. Moreira, "Map-Matching Algorithms for Robot Self-Localization: A Comparison Between Perfect Match, Iterative Closest Point and Normal Distributions Transform," *Journal of Intelligent & Robotic Systems*, vol. 93, no. 3-4, pp. 533–546, mar 2019.

[29] A. Pishehvari, U. Iurgel, S. Lessmann, L. Roese-Koerner, and B. Tibken, "Radar Scan Matching Using Navigation Maps," in *2019 Third IEEE International Conference on Robotic Computing (IRC)*. IEEE, feb 2019, pp. 204–211.

[30] W. Qingshan and Z. Jun, "Point Cloud Registration Algorithm Based on Combination of NDT and PLICP," in *2019 15th International Conference on Computational Intelligence and Security (CIS)*. IEEE, dec 2019, pp. 132–136.

[31] Q.-H. Pham, N.-H. Tran, T.-T. Nguyen, and T.-P. Tran, "Online Robust Sliding-Windowed LiDAR SLAM in Natural Environments," in *2021 International Symposium on Electrical and Electronics Engineering (ISEE)*. IEEE, apr 2021, pp. 172–177.

[32] "VELODYNE sensors' datasheet." [Online]. Available: https://visimind.com/wp-content/uploads/2018/12/LiDAR-Product-Brochure.pdf

[33] "RPLIDAR A2M8 datasheet." [Online]. Available: https://cdn.sparkfun.com/assets/e/a/f/9/8/LD208_SLAMTEC_rplidar_datasheet_A2M8_v1.0_en.pdf

[34] "YDLIDAR sensors' datasheets." [Online]. Available: https://www.manualshelf.com/brand/ydlidar

[35] "Link to repositories for implementation of PLICP, NDT, FastGICP, FastVGICP, and NDT-PSO." [Online]. Available: https://gist.github.com/li9i/c159e282ce5a5ef34500b26c34499c08