Scan-matching of panoramic 2D range scans

Alexandros Filotheou, Antonis Dimitriou, Georgios Sergiadis

***Abstract*— A method for matching coplanar 2D range scans extracted from a LIDAR sensor whose field of view is 2π is proposed. The method exploits properties of the Fourier transform that arise due to the periodicity of the range signal. Matching is performed in a correspondenceless manner. The afforded robustness and the increased accuracy compared to established scan-matcing methods is exhibited through experiments. The source code is available for download.**

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

I. Introduction

Scan-matching is employed in robotics as a means to odometry, primarily in non-wheeled robots where no encoders can be utilised, or as a useful ameliorator of the ever-drifting encoder-ed odometry: scans captured at consecutive time instances, inputted to a scan-matching algorithm, convey an estimate as to the pose of the scan sensor at the second capture time relative to that captured first. Scan-matching is being successfully employed in the tasks of simultaneous localisation and mapping [12]-[14], local map construction [15]-[17], and in people-tracking systems [18].

cannot stress enough: the contributions must be layed out and emphasised see. olsen 2009 for a pro

Contribution: invariant to large angular errors, pseudoinvariant in time of execution, show how to decrease orientation error from fixed number of rays, fourier for both orientation and translation –¿ robustness

II. Problem Formulation

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 $t_1 \geq t_0$

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 $\mathcal{S}_1$.

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

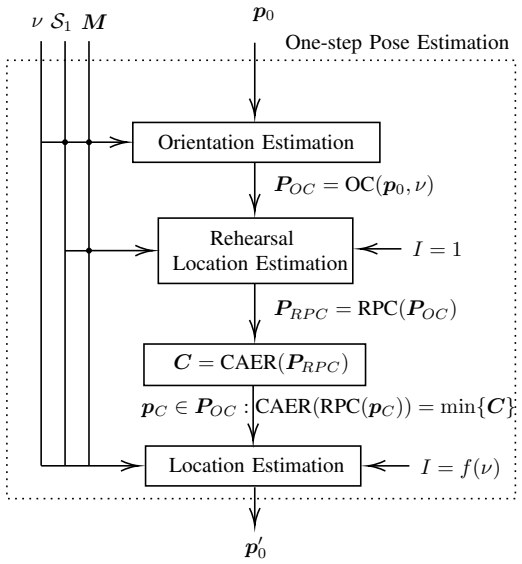

Fig. 1: FSM iteratively invokes the One-step Pose Estimation method. Given a pose estimate of where scan $\mathcal{S}_1$ was captured within $\boldsymbol{M}$, the method attempts to register $\mathcal{S}_1$ to $\boldsymbol{M}$ by estimating first its relative orientation and then its location with respect to the input pose estimate

III. Definitions

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 [11]. We define a range scan $\mathcal{S}$, consisting of N_s rays over an angular range λ, to be an ordered map $\mathcal{S} : \Theta \to \mathbb{R}_{\geq 0}$, where $\Theta = \{\theta_n \in [-\frac{\lambda}{2}, +\frac{\lambda}{2}) : \theta_n = -\frac{\lambda}{2} + \lambda\frac{n}{N_s},\ n = 0, 1, \ldots, N_s - 1\}$. Angles θ_n are expressed relative to the sensor's heading, in the sensor's frame of reference. A LIDAR sensor's angle increment γ is the angular distance between two consecutive rays: $\gamma \triangleq \dfrac{\lambda}{N_s}$.

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 obstacles within a point-set, usually referred to as the map, rather than within an real environment. A map-scan is derived by means of locating intersections of rays emanating from the estimate of the sensor's pose and the boundaries of the map.

IV. Prior Work

V. Approach

The problem 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. The second is estimating the relative displacement of $\mathcal{S}_1$ with respect to $\mathcal{S}_0$ under the assumption that both are captured from poses of the same orientation. Solving the first sub-problem is followed by the solution to the second sub-problem. This process is iterated until termination conditions are met.

The orientation and location estimation submethods are presented in subsections V-A and V-B respectively. Subsection V-C presents the method of how these two are woven together into the system that solves Problem I that is proposed in this work.

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 of signal $\mathcal{S}$, with $\mathcal{F}^{-1}\{\mathcal{S}\}$ its inverse, and with $\boldsymbol{c}^*$ the conjugate of complex $\boldsymbol{c}$, 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 $k, \xi \in [0, N_s - 1]$, $k, \xi \in \mathbb{Z}$:

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

and therefore since $2\pi\frac{\xi}{N_s} = \xi\frac{2\pi}{N_s} = \xi\gamma$

$$\begin{aligned} Q_{\mathcal{S}_0,\mathcal{S}_1}(u) &= \frac{\mathcal{F}\{\mathcal{S}_0\}^* \cdot \mathcal{F}\{\mathcal{S}_1\}}{|\mathcal{F}\{\mathcal{S}_0\}| \cdot |\mathcal{F}\{\mathcal{S}_1\}|} \\ &\simeq \frac{e^{-j\xi\gamma u} \cdot \mathcal{F}\{\mathcal{S}_1\}^* \cdot \mathcal{F}\{\mathcal{S}_1\}}{|e^{-j\xi\gamma u} \cdot \mathcal{F}\{\mathcal{S}_1\}^*| \cdot |\mathcal{F}\{\mathcal{S}_1\}|} \\ &= e^{-j\xi\gamma u} \cdot \frac{\mathcal{F}\{\mathcal{S}_1\}^* \cdot \mathcal{F}\{\mathcal{S}_1\}}{|\mathcal{F}\{\mathcal{S}_1\}| \cdot |\mathcal{F}\{\mathcal{S}_1\}|} \\ &= e^{-j\xi\gamma u} \end{aligned} \tag{2}$$

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 ξ:

$$\xi = \arg\max_u \; q_{\mathcal{S}_0,\mathcal{S}_1}(u) \tag{3}$$

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 twofold, as its existence is also propagated to the location estimation method. In the following, a method for further reduction of the orientation error is presented.

Let $\mathcal{S}_0$ be projected onto the 2D plane around an arbitrary but fixed pose $\boldsymbol{p}_0(x_0, y_0, \theta_0)$, producing point-set $\boldsymbol{M}_R$, which 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_0, 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 $\theta_0^k = \theta_0 + k \cdot \gamma/2^\nu$, 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

$$\mathrm{PD}_k = \frac{2 \max q_{\mathcal{S}_0^k,\mathcal{S}_1}}{\max q_{\mathcal{S}_0^k,\mathcal{S}_0^k} + \max q_{\mathcal{S}_1,\mathcal{S}_1}} \tag{4}$$

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 [23].

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 $\mathcal{S}_V^K$ should be rotated counter-clockwise in order to achieve PD_K: $\Xi = \arg\max \; q_{\mathcal{S}_0^K,\mathcal{S}_1}$. Then the sensor's orientation difference becomes $\Delta\theta = \Xi\gamma + K \cdot \gamma/2^\nu + \delta\theta'$. ?? Consider CAER here ??

If map-scans $\mathcal{S}_V^k$ 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 the guarantee of $|\delta\theta'| \leq \gamma/2^{1+\nu}$ may not always hold for all combinations of environments and sensor angle increments.

B. Estimation of Relative Location

Let the assumptions of Problem I be standing. 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 again onto the 2D plane around fixed pose $\boldsymbol{p}_0(x_0, y_0, \theta_0)$, 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 [24]. 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 $\boldsymbol{M}_L$. Theorem II 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, if the location component of

the arbitrary initial pose $\boldsymbol{p}_0$ is treated according to Theorem I.

Theorem I. *Let a 2D range scan $\mathcal{S}_R$ be captured from a physical panoramic range sensor from unknown pose $\boldsymbol{p} = (\boldsymbol{l}, \theta)$, $\boldsymbol{l} = (x, y)$. Let $\boldsymbol{M}$ be the map of the world in which the scan was captured. Let a pose estimate $\hat{\boldsymbol{p}} = (\hat{\boldsymbol{l}}, \hat{\theta})$ reside in the neighbourhood of $\mathbf{p}$ in the map's frame of reference. Additionally, let $\hat{\theta} = \theta$. Assume that $\mathcal{S}_R$ is disturbance-free, and that the map of the environment captures the latter perfectly. Then, treating the estimate of the location of the sensor as a state variable $\hat{\boldsymbol{l}}[k] = [\hat{x}[k], \hat{y}[k]]^\top$ and updating it according to the difference equation*

$$\hat{\boldsymbol{l}}[k+1] = \hat{\boldsymbol{l}}[k] + \boldsymbol{u}[k] \tag{5}$$

where $\hat{\boldsymbol{l}}[0] = \hat{\boldsymbol{l}} = [\hat{x}, \hat{y}]^\top$, i.e. the supplied initial location estimate, where

$$\boldsymbol{u}[k] = \frac{1}{N_s} \begin{bmatrix} \cos\hat{\theta} & \sin\hat{\theta} \\ \sin\hat{\theta} & -\cos\hat{\theta} \end{bmatrix} \begin{bmatrix} X_{1,r}\big(\mathcal{S}_R, \mathcal{S}_V|_{\hat{\boldsymbol{p}}[k]}\big) \\ X_{1,i}\big(\mathcal{S}_R, \mathcal{S}_V|_{\hat{\boldsymbol{p}}[k]}\big) \end{bmatrix} \tag{6}$$

is the two-dimensional vector hereafter referred to as the control vector, with $X_{1,r}(\cdot)$ and $X_{1,i}(\cdot)$ being, respectively, the real and imaginary parts of the complex quantity X_1:

$$\begin{aligned} X_1\big(\mathcal{S}_R, \mathcal{S}_V|_{\hat{\boldsymbol{p}}[k]}\big) =& X_{1,r}\big(\mathcal{S}_R, \mathcal{S}_V|_{\hat{\boldsymbol{p}}[k]}\big) \\ &+ i \cdot X_{1,i}\big(\mathcal{S}_R, \mathcal{S}_V|_{\hat{\boldsymbol{p}}[k]}\big) \\ =& \sum_{n=0}^{N_s-1} \big(\mathcal{S}_R[n] - \mathcal{S}_V[n]|_{\hat{\boldsymbol{p}}[k]}\big) \cdot e^{-i\frac{2\pi n}{N_s}} \end{aligned} \tag{7}$$

where $\mathcal{S}_R[n]$ and $\mathcal{S}_V[n]|_{\hat{\boldsymbol{p}}[k]}$ are, respectively, the ranges of the n-th ray of real scan $\mathcal{S}_R$ and map-scan $\mathcal{S}_V|_{\hat{\boldsymbol{p}}[k]}$ captured via raycasting the map $\boldsymbol{M}$ from $\hat{\mathbf{p}}[k] = (\hat{\boldsymbol{l}}[k], \hat{\theta})$—then $\hat{\boldsymbol{l}}[k]$ converges to $\boldsymbol{l}$ uniformly asymptotically as $k \to \infty$.

Theorem II. *Let the assumptions of Theorem I hold. Assume additionally that the ranges of either or both real and virtual range scans $\mathcal{S}_R$ and $\mathcal{S}_V$ are affected by additive, bounded disturbances. Then $\hat{\boldsymbol{l}}[k]$ is uniformly bounded for $k \geq k_0$ and uniformly ultimately bounded in a neighbourhood of $\boldsymbol{l}$. Its size depends on the suprema of the disturbance corrupting the range measurements of the two scans.*

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 the point-set that is the result of the projection of range scan $\mathcal{S}_0$ to the $x-y$ plane around an arbitrary but fixed pose $\boldsymbol{p}_0$ by $\boldsymbol{M}$. Then the objective is to estimate the pose $\boldsymbol{p}_1$ from where $\mathcal{S}_1$ was captured relative to $\boldsymbol{p}_0$ by way of registering $\mathcal{S}_1$ to map $\boldsymbol{M}$. In the following, pose $\boldsymbol{p}_0$ is treated as the pose estimate of $\boldsymbol{p}_1$.

Given an input pose estimate $\boldsymbol{p}_0(x_0, y_0, \theta_0)$, range scan $\mathcal{S}_1$, the map $\boldsymbol{M}$, and a sampling degree ν, the One-step Pose Estimation system (fig. 1) first calculates 2^ν pose estimates of $\boldsymbol{p}_1$: $\boldsymbol{P}_{OC} = \{(x_0, y_0, \theta_0^k)\}$, $k = 0, \ldots, 2^\nu - 1$, according to the orientation estimation method described in section V-A.

If the location of $\boldsymbol{p}_0$ coincided with the location of $\boldsymbol{p}_1$, the Percent Discrimination metric (eq. (4)) 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 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 V-B. This operation produces the pose set $\boldsymbol{P}_{RPC} = \{(x_0^k, y_0^k, \theta_0^k)\}$, $|\boldsymbol{P}_{RPC}| = 2^\nu$. 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. This divergence is captured by the Cumulative Absolute Error per Ray (CAER) metric:

$$\text{CAER}_k = \sum_{n=0}^{N_s-1} \left| \mathcal{S}_1[n] - \mathcal{S}_V[n] \Big|_{(\hat{x}_k, \hat{y}_k, \hat{\theta}_k)} \right| \tag{8}$$

where $\mathcal{S}_V$ is the map-scan captured from $(\hat{x}_k, \hat{y}_k, \hat{\theta}_k)$, $k = 0, \ldots, 2^\nu - 1$, within $\boldsymbol{M}$. The CAER metric 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.[1] The pose estimate $\boldsymbol{p}_C \in \boldsymbol{P}_{OC}$ that, 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 system produces $\boldsymbol{p}_0'$, which is then fed back to the Orientation Estimation system in the form of a new pose estimate of $\boldsymbol{p}_1$: $\boldsymbol{p}_0 \leftarrow \boldsymbol{p}_0'$. In practice, the pose set $\boldsymbol{P}_{OC}$ is supplemented with one pose whose location component is equal to $\boldsymbol{p}_0$ 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 $\boldsymbol{p}_0$, range scan $\mathcal{S}_1$, and the map $\boldsymbol{M}$, the pose estimation method proposed iteratively invokes the

[1] Alternatively, correcting the position of 2^ν pose estimates and feeding them back to the One-step Pose Estimation method would incur exponential costs in time of execution.

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 $\boldsymbol{p}_0$ is processed by the One-step Pose Estimation process, and its output $\boldsymbol{p}'_0$ 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, and the process is reset. If no significant pose estimate correction is observed $\|\boldsymbol{p}'_0 - \boldsymbol{p}_0\|_2 < \varepsilon_{\delta p}$, 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 $\boldsymbol{p}'_0$, which is the pose estimate of $\boldsymbol{p}_1$ in the frame of reference of $\boldsymbol{M}$. The roto-translation $\hat{\boldsymbol{T}} = \boldsymbol{p}'_0 - \boldsymbol{p}_0$ is the estimate of the sensor's true motion $\boldsymbol{T}$.

VI. Results

A. Experimental Design

This section serves to test the efficacy and performance of the proposed method. The experimental procedure was conducted using a benchmark dataset D consisting of $|D| = 778$ laser scans obtained from a Sick range-scan sensor mounted on a robotic wheel-chair [25]. For each scan D^d, $d = 1, \ldots, 778$, the dataset reports one range scan of 360 range measurements and the pose from which it was captured $\boldsymbol{r}^d(x, y, \theta)$. The same dataset was used to evaluate the performance of IDC [21], ICP, and MBICP in [3], and that of PLICP and the joint method PLICP∘GPM during scan-matching experiments. In [4] the latter was found to be the best-performing among the five correspondence-finding scan-matching methods. Therefore, for purposes of comparison against correspondence-finding scan-matching methods, the experimental procedure is extended to PLICP∘GPM. This method shall be denoted hereafter by the acronym CSM. In the same vein, for purposes of comparison against correspondenceless scan-matching methods, the same experimental procedure is extended to the Normal Distributions Transform (NDT) scan-matching method [10].

The experimental setup is the following. The rays of each dataset instance D^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. Its radius is set to the minimum range between the two extreme rays of D^d. Similar 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 $U_{xy}(-\overline{\delta}_{xy}, \overline{\delta}_{xy})$, $U_\theta(-\overline{\delta}_\theta, \overline{\delta}_\theta)$; $\overline{\delta}_{xy}$, $\overline{\delta}_\theta \in \mathbb{R}_{\geq 0}$. ?? TODO say that censi et al did not test this but tested self-scan-matching ??

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 CSM, NDT, and FSM are then set to $\mathcal{S}_0^d$ and $\mathcal{S}_1^d$. Their output is $\boldsymbol{p}_0'^d$, which is the pose estimate of $\boldsymbol{p}_1^d$ in the frame of reference of $\boldsymbol{M}^d$. The roto-translation $\hat{\boldsymbol{T}}^d = \boldsymbol{p}_0'^d - \boldsymbol{p}_0^d$ is the estimate of the motion of the range sensor $\boldsymbol{T}^d = \boldsymbol{p}_1^d - \boldsymbol{p}_0^d$. The criterion on which the evaluation of all experiments rests is the 2-norm of the total pose displacement error

$$e = (\Delta \boldsymbol{T}^\top \cdot \Delta \boldsymbol{T})^{1/2} \tag{9}$$

where $\Delta \boldsymbol{T} = \boldsymbol{T} - \hat{\boldsymbol{T}}$; $\boldsymbol{T} = (\Delta x, \Delta y, \Delta\theta) = \boldsymbol{p}_1 - \boldsymbol{p}_0$, and $\hat{\boldsymbol{T}} = (\Delta\hat{x}, \Delta\hat{y}, \Delta\hat{\theta}) = \boldsymbol{p}'_0 - \boldsymbol{p}_0$, For every pose estimate $\boldsymbol{p}_0'^d$ outputted by each algorithm, $d = 1, 2, \ldots, |D|$, its offset from the actual pose $\boldsymbol{p}_1^d$ is recorded in the form of the 2-norm total error. The pose errors of one experiment are then averaged. The pose error distributions reported below are those of mean errors across E experiments of the same configuration.

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 $\sigma_R \in \{0.0, 0.01, 0.03, 0.05, 0.10\}$ m. The non-zero 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 $\sigma_R = 0.01$ m is reported for VELODYNE sensors [26]; the rest are reported for price-appealing but disturbance-laden sensors, e.g. the RPLIDAR A2M8, or the YDLIDAR G4, TG30, and X4 scanners [27]-[30]. 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 $(\mu_{\min}, \mu_{\max}) = (2^{\nu_{\min}}, 2^{\nu_{\max}}) = (2^0, 2^3)$. The number of iterations of the translational component at each map sampling degree ν was set at $I = 2\nu$. The orientation convergence threshold was set to $\varepsilon_{\delta p} = 10\text{e-}5$. Maximal displacements $\overline{\delta}_{xy}$ and $\overline{\delta}_\theta$ were chosen as such by prior art tests [4]. For each experiment FSM, CSM, and NDT ran for $E = 100$ times across all instances of D. All experiments and algorithms were run serially, in C++, on a single thread, on a machine with a CPU frequency of 4.0 GHz. The implementations of CSM and NDT were taken from [31] and [32].

B. Performance

Figure 2 shows the distribution of roto-translation errors (eq. (9)) across E experiments for CSM (red), NDT (blue), and the proposed method of FSM (green).

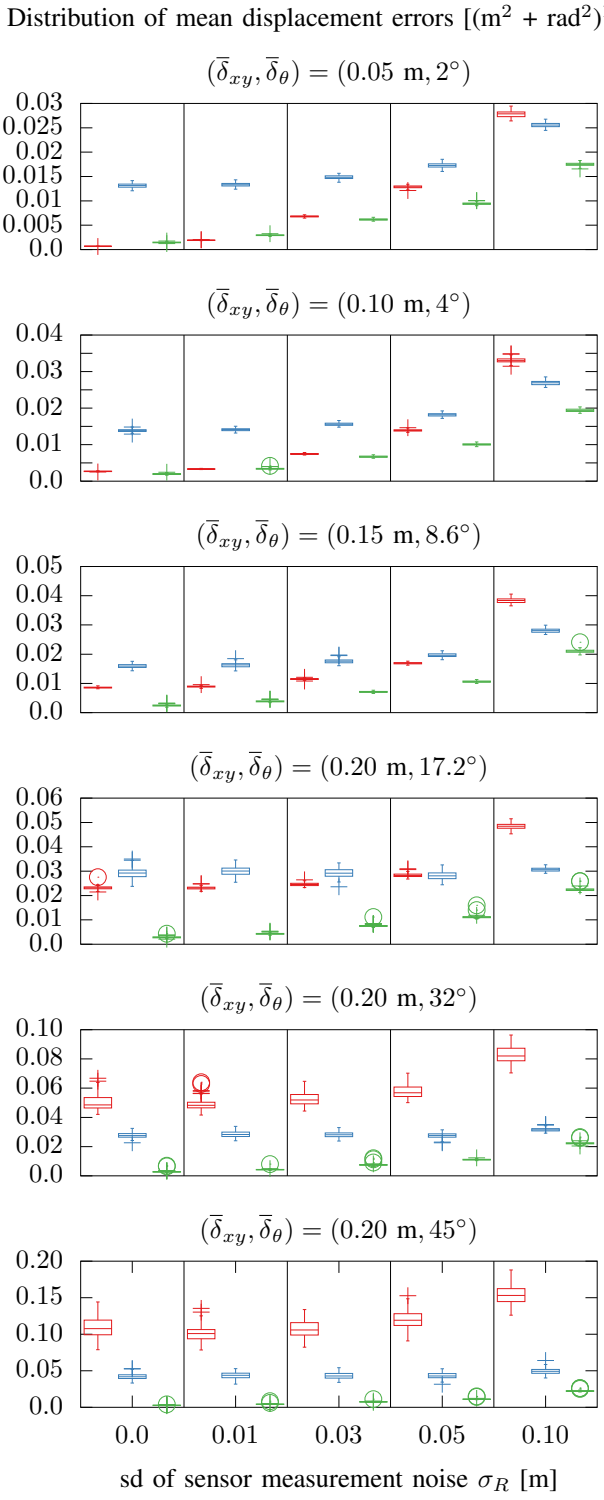

Fig. 2: Distribution of mean errors of CSM (red), NDT (blue), and FSM (our approach) across a range of maximal positional and orientational displacements, for progressively larger sensor measurement noise levels. The variability of FSM's rigid body transformation error is consistent across all configurations. The error is independent of the initial displacement of scans for a given level of sensor noise

At small location and orientation displacements between the two input scans ($\overline{\delta}_{xy} \leq 0.05$ m, $\overline{\delta}_\theta \leq 2°$), CSM outperforms NDT and FSM for low levels of sensor noise ($\sigma_R \leq 0.01$ m). However, as noise increases, FSM starts exhibiting greater robustness and accuracy than CSM. At greater location and orientation displacements ($\overline{\delta}_{xy} > 0.05$ m, $\overline{\delta}_\theta > 2°$), FSM is able to maintain errors equal to or lower than CSM across the entirety of the spectrum of tested noise levels. Compared to the equally correspondenceless method of NDT, FSM exhibits greater accuracy across all tested configurations. The variability of FSM's errors for a given level of sensor noise is independent of the displacement of the two input scans. The juxtaposition of the three methods' pose 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, CSM ranged from 4.8 to 20.5 ms, NDT from 8.1 to 19.9 ms, and FSM from 13.2 to 16.7 ms. Therefore FSM's exhibits the least variability to sensor noise and locational and orientational displacement in terms of runtime. The measurement frequency of modern LIDAR sensors ranges from 12-20 Hz; therefore FSM runs in real time in modern processors.

VII. Conclusions

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. The C++ code of the proposed method, along with the implementation of the conducted experiments is available at , .

References

[1] Sebastian Thrun, Wolfram Burgard, and Dieter Fox, "Probabilistic Robotics" (Intelligent Robotics and Autonomous Agents), The MIT Press, 2005

[2] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes", IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, volume 14, number 2, pp. 239-256, doi 10.1109/34.121791, ISSN 0162-8828

[3] J. Minguez, F. Lamiraux and L. Montesano, "Metric-Based Scan Matching Algorithms for Mobile Robot Displacement Estimation," Proceedings of the 2005 IEEE International Conference on Robotics and Automation, Barcelona, Spain, 2005, pp. 3557-3563, doi: 10.1109/ROBOT.2005.1570661

[4] A. Censi, "An ICP variant using a point-to-line metric," 2008 IEEE International Conference on Robotics and Automation, Pasadena, CA, 2008, pp. 19-25, doi: 10.1109/ROBOT.2008.4543181.

[5] Wang, J. Zhao, M. Chen, W. "MIM-SLAM: A Multi-Level ICP Matching Method for Mobile Robot in Large-Scale and Sparse Scenes". Appl. Sci. 2018, 8, 2432.

[6] Tian, Yingzhong; Liu, Xining; Li, Long; Wang, Wenbin. 2019. "Intensity-Assisted ICP for Fast Registration of 2D-LIDAR" Sensors 19, no. 9: 2124.

[7] Naus, Krzysztof; Marchel, ukasz. 2019. "Use of a Weighted ICP Algorithm to Precisely Determine USV Movement Parameters" Appl. Sci. 9, no. 17: 3530.

[8] Linh Tao, Tam Bui, Toshio Ito, "MODIFIED ITERATIVE CLOSEST POINT MATCHING FOR 2D LIDAR LASER DATA", 14th South East Asian Technical University Consortium 2020 (SEATUC 2020) 27th 28th February 2020, KX Building, KMUTT, Bangkok, Thailand

[9] Huchao Xu, Letao Zhou, Yinghao Zhao, Zheng Yuan, " A Two-Dimensional Point Cloud Matching Method Based on ICP Improvement", China Satellite Navigation Conference (CSNC) 2020 Proceedings: Volume I pp 390-398

[10] P. Biber and W. Strasser, "The normal distributions transform: a new approach to laser scan matching," Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453), Las Vegas, NV, USA, 2003, pp. 2743-2748 vol.3, doi: 10.1109/IROS.2003.1249285.

[11] Cooper, M.A.; Raquet, J.F.; Patton, R. Range Information Characterization of the Hokuyo UST-20LX LIDAR Sensor. Photonics 2018, 5, 12.

[12] J. -. Gutmann and K. Konolige, "Incremental mapping of large cyclic environments" Proceedings 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation. CIRA'99 (Cat. No.99EX375), Monterey, CA, USA, 1999, pp. 318-325. doi: 10.1109/CIRA.1999.810068

[13] D. Hahnel, W. Burgard, D. Fox and S. Thrun, "An efficient fastSLAM algorithm for generating maps of large-scale cyclic environments from raw laser range measurements," Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453), Las Vegas, NV, USA, 2003, pp. 206-211 vol.1. doi: 10.1109/IROS.2003.1250629

[14] Chieh-Chih Wang, C. Thorpe and S. Thrun, "Online simultaneous localization and mapping with detection and tracking of moving objects: theory and results from a ground vehicle in crowded urban areas," 2003 IEEE International Conference on Robotics and Automation (Cat. No.03CH37422), Taipei, Taiwan, 2003, pp. 842-849 vol.1. doi: 10.1109/ROBOT.2003.1241698

[15] Lacroix, S., Mallet, A., Bonnafous, D., Bauzil, G., Fleury, S., Herrb, M., and Chatila, R. (2002). "Autonomous Rover Navigation on Unknown Terrains: Functions and Integration". The International Journal of Robotics Research, 21(1011), 917942. https://doi.org/10.1177/027836490202101084 1

[16] Minguez, J., Montesano, L., anc Montano, L. (2004). "An architecture for sensor-based navigation in realistic dynamic and troublesome scenarios". 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566), 3, 2750-2756 vol.3.

[17] Montesano, Luis, Minguez, Javier and Montano, Luis, "Modeling dynamic scenarios for local sensor-based motion planning", Autonomous Robots, 2008, Volume 25, pp 231-251.

[18] D. Schulz, W. Burgard, D. Fox and A. B. Cremers, "Tracking multiple moving targets with a mobile robot using particle filters and statistical data association", Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No.01CH37164), Seoul, South Korea, 2001, pp. 1665-1670 vol.2. doi: 10.1109/ROBOT.2001.932850

[19] M. Leordeanu and M. Hebert, "A spectral technique for correspondence problems using pairwise constraints", Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, Beijing, 2005, pp. 1482-1489 Vol. 2. doi: 10.1109/ICCV.2005.20

[20] Filotheou, A., Tsardoulias, E., Dimitriou, A. et al. "Pose Selection and Feedback Methods in Tandem Combinations of Particle Filters with Scan-Matching for 2D Mobile Robot Localisation". J Intell Robot Syst 100, 925944 (2020). https://doi.org/10.1007/s10846-020-01253-6

[21] Feng Lu and Milios, "Robot pose estimation in unknown environments by matching 2D range scans," 1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 1994, pp. 935-938.

[22] A. Censi, "Scan matching in a probabilistic framework," Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006., Orlando, FL, 2006, pp. 2291-2296, doi: 10.1109/ROBOT.2006.1642044.

[23] Qin-Sheng Chen, M. Defrise and F. Deconinck, "Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 12, pp. 1156-1168, Dec. 1994, doi: 10.1109/34.387491.

[24] E. B. Olson, "Real-time correlative scan matching," 2009 IEEE International Conference on Robotics and Automation, Kobe, Japan, 2009, pp. 4387-4393, doi: 10.1109/ROBOT.2009.5152375.

[25] The dataset is available at `https://censi.science/pub/research/2007-plicp/laserazosSM3.log.gz`, last accessed 24 Mar 2021

[26] VELODYNE sensors' datasheet `https://visimind.com/wp-content/uploads/2018/12/LiDAR-Product-Brochure.pdf`, last accessed 24 Mar 2021

[27] RPLIDAR A2M8 datasheet `https://cdn.sparkfun.com/assets/e/a/f/9/8/LD208_SLAMTEC_rplidar_datasheet_A2M8_v1.0_en.pdf`, last accessed 24 Mar 2021

[28] YDLIDAR G4 datasheet `https://www.ydlidar.com/Public/upload/files/2020-10-29/YDLIDAR%20G4%20Datasheet.pdf`, last accessed 24 Mar 2021

[29] YDLIDAR TG30 datasheet `https://www.ydlidar.com/Public/upload/files/2020-10-13/YDLIDAR%20TG30%20Data%20sheet.pdf`, last accessed 24 Mar 2021

[30] YDLIDAR X4 datasheet `https://www.ydlidar.com/Public/upload/files/2020-04-13/YDLIDAR%20X4%20Datasheet.pdf`, last accessed 24 Mar 2021

[31] `https://github.com/AndreaCensi/csm`, last accessed 24 Mar 2021

[32] `https://github.com/PointCloudLibrary/pcl/blob/master/registration/include/pcl/registration/impl/ndt.hpp`, last accessed 24 Mar 2021

[33] F.A. Donoso, K.J. Austin, P.R. McAree, "How do ICP variants perform when used for scan matching terrain point clouds?," Robotics and Autonomous Systems, Volume 87, 2017, Pages 147-161, ISSN 0921-8890, https://doi.org/10.1016/j.robot.2016.10.011.