







KISS-ICP: In Defense of Point-to-Point ICP – Simple, Accurate, and Robust Registration If Done the Right Way

Ignacio Vizzo , *Student Member, IEEE*, Tiziano Guadagnino , *Member, IEEE*,
Benedikt Mersch , *Graduate Student Member, IEEE*, Louis Wiesmann , Jens Behley , *Member, IEEE*,
and Cyrill Stachniss , *Member, IEEE*

Abstract—Robust and accurate pose estimation of a robotic platform, so-called sensor-based odometry, is an essential part of many robotic applications. While many sensor odometry systems made progress by adding more complexity to the ego-motion estimation process, we move in the opposite direction. By removing a majority of parts and focusing on the core elements, we obtain a surprisingly effective system that is simple to realize and can operate under various environmental conditions using different LiDAR sensors. Our odometry estimation approach relies on point-to-point ICP combined with adaptive thresholding for correspondence matching, a robust kernel, a simple but widely applicable motion compensation approach, and a point cloud subsampling strategy. This yields a system with only a few parameters that in most cases do not even have to be tuned to a specific LiDAR sensor. Our system performs on par with state-of-the-art methods under various operating conditions using different platforms using the same parameters: automotive platforms, UAV-based operation, vehicles like segways, or handheld LiDARs. We do not require integrating IMU data and solely rely on 3D point clouds obtained from a wide range of 3D LiDAR sensors, thus, enabling a broad spectrum of different applications and operating conditions. Our open-source system operates faster than the sensor frame rate in all presented datasets and is designed for real-world scenarios.

Index Terms—Mapping, localization, SLAM.

I. INTRODUCTION

ODOMETRY estimation is an essential building block for any mobile robot that needs to autonomously navigate in unknown environments. In the LiDAR sensing domain, current odometry pipelines typically use some form of (ICP) to estimate poses incrementally [10], [26], [31], [35]. Even though LiDAR odometry has been an active area of research for the last three

decades, the design of current systems is usually coupled with assumptions about the robot motion [10] and the structure of the environment [28] to achieve accurate and robust alignment results. To the best of our knowledge, no existing 3D LiDAR odometry approach is free of parameter tuning and works out of the box in different scenarios, using arbitrary LiDAR sensors, supporting different motion profiles, and consequently types of robots, such as ground and aerial robots.

This paper returns to the roots: classical point-to-point ICP, introduced 30 years ago by Besl and McKay [3]. We aim to tackle the inherent problems of sequentially operating LiDAR odometry systems that prohibit current approaches from generalizing to different environments, sensor resolutions, and motion profiles using a single configuration. We present simple yet effective reasoning about the robot kinematics and the sequential way LiDAR data is recorded on a mobile platform, as well as an effective downsampled point cloud representation that allows us to minimize the need for parameter tuning.

Our system challenges even extensively hand-tuned and optimized existing simultaneous localization and mapping (SLAM) systems. Our design uses neither sophisticated feature extraction techniques, learning methods, nor loop closures. The same parameter set works in various challenging scenarios such as highway drives of robot cars with many dynamic objects, drone flights, handheld devices, segways, and more. Thus, we take a step back from mainstream research in LiDAR odometry estimation and focus on reducing the components to their essentials. This makes our system perform extraordinarily well in various real-world scenarios, see Fig. 1.

The main contribution of this paper is a simple yet highly effective approach for building LiDAR odometry systems that can accurately compute a robot's pose online while navigating through an environment. We identify the core components and properly evaluate the impact of different modules on such systems. We show that with the proper use of ICP that builds on basic reasoning about the system's physics and the sensor data's nature, we obtain competitive odometry. Besides motion prediction, spatial scan downsampling, and a robust kernel, we introduce an adaptive threshold approach for ICP in the context of robot motion estimation that makes our approach effective and, at the same time, generalizes easily.

We make three key claims: Our “keep it small and simple” approach exploiting point-to-point ICP is (i) on par with state-of-the-art odometry systems, (ii) can accurately compute the robot's odometry in a large variety of environments and motion

Manuscript received 14 September 2022; accepted 27 December 2022. Date of publication 12 January 2023; date of current version 18 January 2023. This letter was recommended for publication by Associate Editor J. L. Blanco-Claraco and Editor J. Civera upon evaluation of the reviewers comments. This work was supported in part by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy under Grant EXC-2070 – 390732324 – PhenoRob and in part by the European Union's HORIZON Research and Innovation Programme under Grant 101070405 (Digiforest). (Corresponding author: Ignacio Vizzo.)

The authors are with the University of Bonn, 53113 Bonn, Germany, also with the Department of Engineering Science, University of Oxford, Oxford OX1 2JD, United Kingdom, and also with the Lamarr Institute for Machine Learning and Artificial Intelligence, IU International University of Applied Sciences, 10247 Berlin, Germany (e-mail: cyrill.stachniss@igg.uni-bonn.de).

Digital Object Identifier 10.1109/LRA.2023.3236571

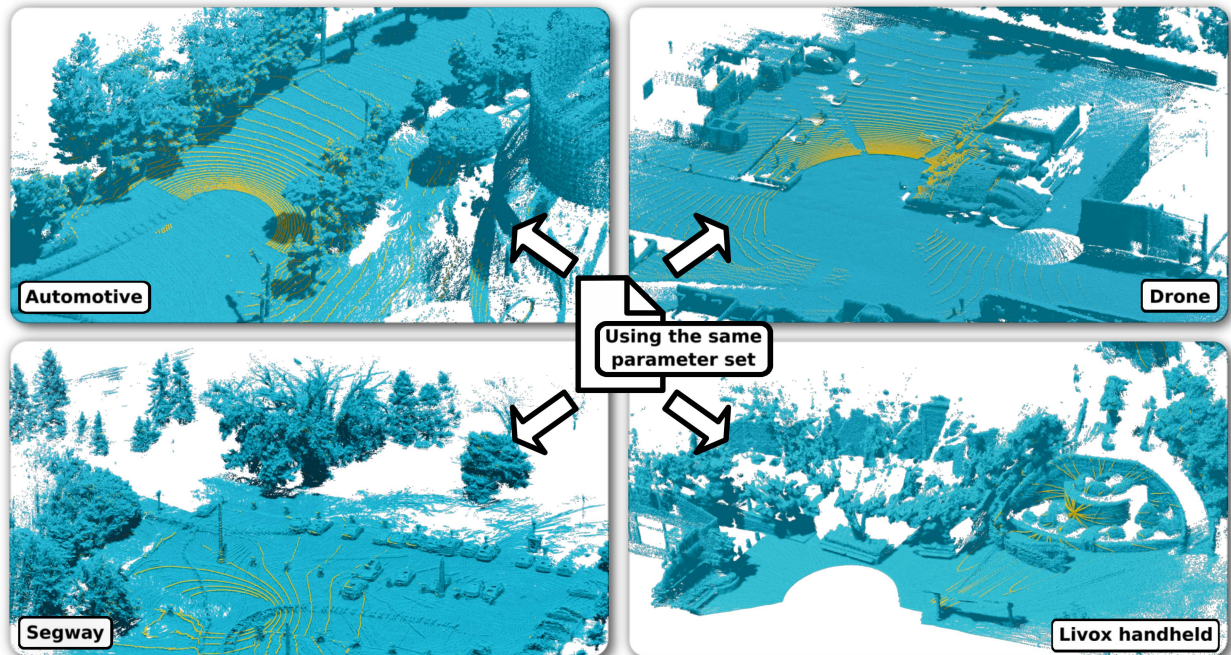


Fig. 1. Point cloud maps (blue) generated by our proposed odometry pipeline on different datasets with the same set of parameters. We depict the latest scan in yellow. The scans are recorded using different sensors with different point densities, different orientations, and different shooting patterns. The automotive example stems from the MulRan dataset [15]. The drone of the Voxgraph dataset [23] and the segway robot used in the NCLT dataset [5] show a high acceleration motion profile. The handheld Livox LiDAR [17] has a completely different shooting pattern than the commonly used rotating mechanical LiDAR.

profiles with the same system configuration, and (iii) provides an effective solution to motion distortion without relying on IMUs or wheel odometers. In sum, “good old point-to-point ICP” is a surprisingly powerful tool, and there is little need to move to more sophisticated approaches if the basic components are done well.

We provide an open-source implementation at: <https://github.com/PRBonn/kiss-icp> that precisely follows the description of this paper.

II. RELATED WORK

Point cloud registration has been an active area of research for the last three decades [3], [9] and is still relevant nowadays. The ICP algorithm can solve the problem of finding a transformation that brings two different point clouds into a common reference frame, and it is a special case of the absolute orientation problem in photogrammetry. ICP typically consists of two parts. The first one is to find correspondences between the point clouds. The second one computes the transformation that minimizes an objective function defined on the correspondences from the first step. One repeats this process until a convergence criterion is met. Most ICP variants [1], [10], [11], [21], [26], [35] utilize a maximum distance threshold in the data association module plus a robust kernel [6] and a maximum number of iterations. In contrast, we propose a threshold estimation method that adapts to changing scenarios by reasoning about the system kinematics and the nature of the data in combination with a robust kernel. We avoid controlling the number of iterations of the ICP to achieve better generalization.

ICP can be used to obtain an odometry estimation from streaming data from a sensor such as RGB-D cameras [19] or

LiDARs [10]. In this work, we focus on the problem of LiDAR odometry estimation, although the ideas presented can be easily extended to other range-sensing technologies.

Nearly all modern SLAM systems build on top of odometry algorithms. Zhang et al. [35] proposed (LOAM) that computes the robot’s odometry by registering planar and edge features to a sparse feature map. LOAM inspired numerous other works [27], [33], such as Lego-LOAM [28], which adds ground constraints to improve accuracy, and recently F-LOAM [33], which revised the original method with a more efficient optimization technique enabling faster operation. However, these methods rely on hand-tuned feature extraction, which typically requires tedious parameter tuning that depends on sensor resolution, environment structure, etc. In contrast, we only rely on point coordinates removing this data-dependent parameter adaptation.

Behley and Stachniss [1] propose the surfel-based method SuMa to achieve LiDAR odometry estimation and mapping. It has also been extended to account for semantics [8] and explicitly handle dynamic objects [7]. In contrast to the surfel-based mapping, Deschaud [11] introduced IMLS-SLAM [11] selecting an implicit moving least square surface [16] as map representation. Along these lines, Vizzo et al. [31] exploited a triangular mesh as the internal map representation. All the above approaches rely on a point-to-plane [24] metric to register consecutive scans. This requires normal estimation, which introduces additional data-dependent parameters. Furthermore, noisy 3D information can impact the normal computation and subsequently the registration in a negative way. We will show that by minimizing a simpler point-to-point metric, we obtain on-par or better odometry performance. Moreover, this design choice enables us to represent the internal map as a voxelized, downsampled point cloud, simplifying the implementation.

Recently, several new approaches [10], [21], [27] have been proposed to solve the odometry estimation problem. Most of these works focus on the runtime operation of the system as well as on the accuracy. Pan et al. [21] propose a multi-metric system (MULLS) that obtains good results in many challenging scenarios at the cost of tuning many parameters for each run. Dellenbach et al. [10] introduced a novel approach, called continuous time ICP (CT-ICP), which incorporates the motion undistortion into the registration showing great results but adding more complexity. Additionally, the robots' motion profile must be known a priori, as, for example, a car will have a different profile than a segway platform. We challenge the need for sophisticated optimization techniques to cope with motion distortion requiring only the constant velocity model. Furthermore, our system only relies on a few parameters, and we do not need to know the motion profile in advance.

Many state-of-the-art odometry systems [1], [10], [21], [27] also rely on pose graph optimization to achieve a better alignment. In contrast, we do not exploit such techniques and state that pose graph optimization is orthogonal to the presented approach and can be easily integrated. In sum, we step back from the common mainstream work on LiDAR odometry and propose a system that solely relies on a point-to-point metric and does not employ pose graph optimization [1], [10], [21], [27]. Our system can run on different types of mobile robots, drones, handheld devices, and segways, without the need to fine-tune the system to a specific application.

III. KISS-ICP – KEEP IT SMALL AND SIMPLE

This work aims to incrementally compute the trajectory of a moving LiDAR sensor by sequentially registering the point clouds recorded by the scanner. We reduced the components to a minimal set needed to build an effective, accurate, robust, and still reasonably simple LiDAR odometry system.

For each 3D scan in form of a local, egocentric point cloud $\mathcal{P} = \{\mathbf{p}_i \mid \mathbf{p}_i \in \mathbb{R}^3\}$, we perform the following four steps to obtain a global pose estimate $\mathbf{T}_t \in SE(3)$ at time t . First, we apply sensor motion prediction and motion compensation, often called deskewing, to undo the distortions of the 3D data caused by the sensor's motion during scanning. Second, we subsample the current scan. Third, we estimate correspondences between the input point cloud and a reference point cloud, which we call the local map. We use an adaptive thresholding scheme for correspondence estimation, restricting possible data associations and filtering out potential outliers. Fourth, we register the input point cloud to the local map using a robust point-to-point ICP algorithm. Finally, we update the local map with a downsampled version of the registered scan. Below, we describe these components in detail.

A. Step 1: Motion Prediction and Scan Deskewing

We advocate for rethinking the point cloud registration in the context of mobile robots, which continuously record data. One should not think of it as registering arbitrary pairs of 3D point clouds. Instead, one should phrase it as *estimating how much the robot's actual motion deviates from its expected motion by registering consecutive scans*.

Different approaches can be used to compute the robot's expected motion before considering the LiDAR data. The three most popular choices are the constant velocity model, wheel

odometry obtained through encoders, and IMU-based motion estimation. The constant velocity [29] model assumes that a robot moves with the same translational and rotational velocity as in the previous time step. It requires no additional sensors (no wheel encoder, no IMU) and thus is the most widely applicable option.

Our approach uses the constant velocity model for two reasons: first, it is generally applicable, requires no additional sensors, and avoids the need for time synchronization between sensors. Second, as we will show in our experimental evaluation, it works well enough to provide a solid initial guess when searching for data associations and deskewing 3D scans. This follows from the fact that robotic LiDAR sensors commonly record and stream point clouds at 10 Hz to 20 Hz, i.e., every 0.05 s to 0.1 s. In most cases, the acceleration or deceleration, i.e., the deviations from the constant velocity model that occurs within such short time intervals, are fairly small. If the robot accelerates or decelerates, the constant velocity estimation of the robot's pose will be slightly off, and therefore, we need to correct this estimate through registration. These accelerations determine the possible displacements of the (static) 3D points.

The constant velocity model approximates the translational and angular velocities, denoted as \mathbf{v}_t and $\boldsymbol{\omega}_t$ at time t respectively, by using the previous pose estimates $\mathbf{T}_{t-1} = (\mathbf{R}_{t-1}, \mathbf{t}_{t-1})$ and $\mathbf{T}_{t-2} = (\mathbf{R}_{t-2}, \mathbf{t}_{t-2})$, represented by a rotation matrix $\mathbf{R}_t \in SO(3)$ and a translation vector $\mathbf{t}_t \in \mathbb{R}^3$ for the time step t . We first compute the relative pose $\mathbf{T}_{\text{pred},t}$ that we will use as motion prediction as:

$$\mathbf{T}_{\text{pred},t} = \begin{bmatrix} \mathbf{R}_{t-2}^\top \mathbf{R}_{t-1} & \mathbf{R}_{t-2}^\top (\mathbf{t}_{t-1} - \mathbf{t}_{t-2}) \\ \mathbf{0} & 1 \end{bmatrix}, \quad (1)$$

then derive the corresponding velocities as:

$$\mathbf{v}_t = \frac{\mathbf{R}_{t-2}^\top (\mathbf{t}_{t-1} - \mathbf{t}_{t-2})}{\Delta t}, \quad (2)$$

$$\boldsymbol{\omega}_t = \frac{\text{Log}(\mathbf{R}_{t-2}^\top \mathbf{R}_{t-1})}{\Delta t}, \quad (3)$$

where Δt is the acquisition time of one LiDAR sweep, typically 0.05 s or 0.1 s, and $\text{Log}: SO(3) \rightarrow \mathbb{R}^3$ extracts the axis-angle representation.

Note that also wheel odometry or an IMU-based motion prediction approach can be used instead to compute \mathbf{v}_t and $\boldsymbol{\omega}_t$ for each time step. This will not change the remainder of our approach. For example, if one has good wheel odometry available, this can also be used. However, we use constant velocity as a generally applicable approach.

Within the acquisition time Δt of one LiDAR sweep, multiple 3D points are measured by the scanner. The relative timestamp $s_i \in [0, \Delta t]$ for each point $\mathbf{p}_i \in \mathcal{P}$ describes the recording time relative to the scan's first measurement. This relative timestamp allows us to compute the motion compensation resulting in a deskewed point $\mathbf{p}_i^* \in \mathcal{P}^*$ of the corrected scan \mathcal{P}^* reading by

$$\mathbf{p}_i^* = \text{Exp}(s_i \boldsymbol{\omega}_t) \mathbf{p}_i + s_i \mathbf{v}_t, \quad (4)$$

where $\text{Exp}: \mathbb{R}^3 \rightarrow SO(3)$ computes a rotation matrix from an axis-angle representation. Note that $\text{Exp}(s_i \boldsymbol{\omega}_t)$ is equivalent to performing SLERP in the axis-angle domain.

This form of scan deskewing, especially with the constant velocity model, is easy to implement, generally applicable,

and does not require additional sensors, high-precision time synchronization between sensors, or IMU biases to be estimated. As we show in Section IV, this approach often performs even better than more complex compensation systems [10], at least as long the motion between the start and end of the sweep is small as it is for most robotics applications.

B. Step 2: Point Cloud Subsampling

Identifying a set of keypoints in the point cloud is a common approach for scan registration [14], [24], [35]. It is typically done to achieve faster convergence and/or higher robustness in the data association. However, complex filtering of the point cloud usually comes with an extra layer of complexity and parameters that often need to be tuned.

Rather than extracting 3D keypoints, which often requires environment-dependent parameter tuning, we propose to compute only a spatially downsampled version $\hat{\mathcal{P}}^*$ of the deskewed scan \mathcal{P}^* . Downsampling is done using a voxel grid. As we will explain in Section III-C below in more detail, we use a voxel grid as our local map, where each voxel cell has a size of $v \times v \times v$ and each cell only store a certain number of points. Every time we process an incoming scan, we first downsample the point cloud of the scan to an intermediate point cloud $\mathcal{P}_{\text{merge}}^*$, which is later used to update the map when the relative motion of the robot has been determined with ICP. To obtain the points in $\mathcal{P}_{\text{merge}}^*$, we use voxel size αv with $\alpha \in (0.0, 1.0]$ and keep only a single point per voxel.

For the ICP registration, an even lower resolution scan is beneficiary. Thus, we compute a further reduced point cloud $\hat{\mathcal{P}}^*$ by downsampling $\mathcal{P}_{\text{merge}}^*$ again using a voxel size of βv with $\beta \in [1.0, 2.0]$ keeping only a single point per voxel. This further reduces the number of points processed during the registration and allows for a fast and highly effective alignment. The idea of this “double downsampling” stems from CT-ICP [10], the so far best performing open-source LiDAR odometry system on KITTI.

Most voxelization approaches, however, select the center of each occupied voxel to downsample the point cloud [25], [36]. Instead, we found it advantageous to maintain the original point coordinates, select only one point per voxel for a single scan, and keep its coordinates to avoid discretization errors. This means the reduced cloud is a subset of the deskewed one, i.e., $\hat{\mathcal{P}}^* \subseteq \mathcal{P}^*$. In our implementation, we keep only the first point that was inserted into the voxel.

C. Step 3: Local Map and Correspondence Estimation

In line with prior work [1], [10], [19], [35], we register the deskewed and subsampled scan $\hat{\mathcal{P}}^*$ to the point cloud built so far, i.e., a local map, to compute an incremental pose estimate $\Delta \mathbf{T}_{\text{icp}}$. We use frame-to-map registration as it proves more reliable and robust than the frame-to-frame alignment [1], [19]. To do that effectively, we must define a data structure representing the previously registered scans.

Modern approaches have used very different types of representations for this local map. Popular approaches are voxel grids [35], triangle meshes [31], surfel representations [1], or implicit representations [11]. As mentioned in Section III-B, we utilize a voxel grid to store a subset of 3D points. We use a grid with a voxel size of $v \times v \times v$ and store up to N_{max} points per

voxel. After registration, we update the voxel grid by adding the points $\{\mathbf{T}_t \mathbf{p} \mid \mathbf{p} \in \mathcal{P}_{\text{merge}}^*\}$ from the new scan using the global pose estimate \mathbf{T}_t . Voxels that already contain N_{max} points are not updated. Additionally, given the current pose estimate, we remove voxels outside the maximum range r_{max} . Thus, the size of the map will stay bounded.

Instead of a 3D array, we use a hash table to store the voxels, allowing a memory-efficient representation and fast nearest neighbor search [10], [20]. However, the used data structure can be easily replaced with VDBs [18], [32], Octrees [30], [34], or KD-Trees [2].

D. Adaptive Threshold for Data Association

ICP typically performs a nearest neighbor data association to find corresponding points between two point clouds [3]. When searching for associations, it is common to impose a maximum distance between corresponding points, often using a value of 1 m or 2 m [1], [31], [35]. This maximum distance threshold can be seen as an outlier rejection scheme, as all correspondences with a distance larger than this threshold are considered outliers and are ignored.

The required value for this threshold τ depends on the expected initial pose error, the number and type of dynamic objects in the scene, and, to some degree, the sensor noise. It is typically selected heuristically. Based on the considerations about the constant velocity motion prediction in Section III-A, we can, however, estimate a likely limit from data by analyzing how much the odometry may deviate from the motion prediction over time. This deviation $\Delta \mathbf{T}$ in the pose corresponds exactly to the local ICP correction to be applied to the predicted pose (but it is not known beforehand). Intuitively, we can observe the robot’s acceleration in the magnitude of $\Delta \mathbf{T}$. If the robot is not accelerating, then $\Delta \mathbf{T}$ will have a small magnitude, often around zero, meaning that the constant velocity assumption holds and no correction has to be done by ICP.

We integrate this information into our data association search by exploiting the so-far successful ICP executions. We can estimate the possible point displacement between corresponding points in successive scans in the presence of a potential acceleration expressed through $\Delta \mathbf{T}$ as:

$$\delta(\Delta \mathbf{T}) = \delta_{\text{rot}}(\Delta \mathbf{R}) + \delta_{\text{trans}}(\Delta \mathbf{t}), \quad (5)$$

where $\Delta \mathbf{R} \in SO(3)$ and $\Delta \mathbf{t} \in \mathbb{R}^3$ refer to the rotational and translational component of the deviation, given by

$$\delta_{\text{rot}}(\Delta \mathbf{R}) = 2 r_{\text{max}} \sin \left(\underbrace{\frac{1}{2} \arccos \left(\frac{\text{tr}(\Delta \mathbf{R}) - 1}{2} \right)}_{\theta} \right) \quad (6)$$

$$\delta_{\text{trans}}(\Delta \mathbf{t}) = \|\Delta \mathbf{t}\|_2. \quad (7)$$

The term $\delta_{\text{rot}}(\Delta \mathbf{R})$ represents the displacement that occurs for a range reading with maximum range r_{max} subject to the rotation $\Delta \mathbf{R}$, see also Fig. 2. Note that (5) constitutes an upper bound for the point displacement as

$$\|\Delta \mathbf{R} \mathbf{p} + \Delta \mathbf{t} - \mathbf{p}\|_2 \leq \delta_{\text{rot}}(\Delta \mathbf{R}) + \delta_{\text{trans}}(\Delta \mathbf{t}), \quad (8)$$

which follows from the triangle inequality.

For obtaining δ_{rot} other approaches could be considered, like taking into account the individual ranges for the adaptive threshold computation [4]. In our tests, we did not see any difference

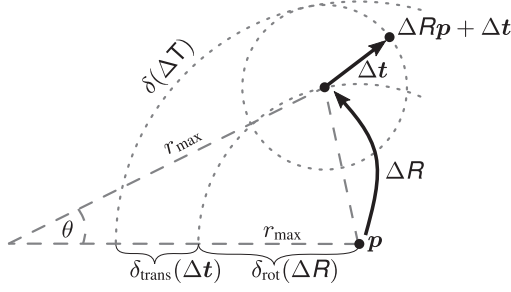


Fig. 2. Exemplary computation of the maximum point displacement $\delta(\Delta T)$ caused by a rotational and translational deviation $(\Delta R, \Delta t)$ from the predicted motion.

in the results but a 3-fold increase of the overall runtime; thus, we use r_{\max} instead of r for computing δ_{rot} .

To compute the threshold τ_t at time t , we consider a Gaussian distribution over δ using the values of (5) over the trajectory computed so far whenever the deviation was larger than a minimum distance δ_{\min} , i.e., situations where the robot's motion was deviating from the constant velocity model. Its standard deviation is

$$\sigma_t = \sqrt{\frac{1}{|\mathcal{M}_t|} \sum_{i \in \mathcal{M}_t} \delta(\Delta T_i)^2}, \quad (9)$$

where the index set \mathcal{M}_t of deviations up to t is given by

$$\mathcal{M}_t = \{i \mid i < t \wedge \delta(\Delta T_i) > \delta_{\min}\}. \quad (10)$$

This avoids reducing the value of σ_t too much when the robot is not moving or is moving at constant velocity for a long time. In our experiments, we set this threshold δ_{\min} to 0.1 m. We then compute the threshold τ_t as the three-sigma bound $\tau_t = 3\sigma_t$, which we use in the next section for the data association search.

E. Step 4: Alignment Through Robust Optimization

We base our registration on classic point-to-point ICP [3]. The advantage of this choice is that we do not need to compute data-dependent features such as normals, curvature, or other descriptors, which may depend on the scanner or the environment. Furthermore, with noisy or sparse LiDAR scanners, features such as normals are often not very reliable. Thus, neglecting quantities such as normals in the alignment process is an explicit design decision that allows our system to generalize well to different sensor resolutions.

To obtain the global estimation of the pose T_t of the robot, we start by applying our prediction model $T_{\text{pred},t}$ to the scan \hat{P}^* in the local frame. Successively, we transform it into the global coordinate frame using the previous pose estimate T_{t-1} , resulting in the source points

$$\mathcal{S} = \{s_i = T_{t-1} T_{\text{pred},t} p \mid p \in \hat{P}^*\}. \quad (11)$$

For each iteration j of ICP, we obtain a set of correspondences between the point cloud \mathcal{S} and the local map $\mathcal{Q} = \{q_i \mid q_i \in \mathbb{R}^3\}$ through nearest neighbor search over the voxel grid (Section III-C) considering only correspondences with a point-to-point distance below τ_t . To compute the current pose correction $\Delta T_{\text{est},j}$, we perform a robust optimization minimizing the sum

TABLE I
ALL SEVEN PARAMETERS OF OUR APPROACH

Parameter	Value
Initial threshold τ_0	2 m
Min. deviation threshold δ_{\min}	0.1 m
Max. points per voxel N_{\max}	20
Voxel size map v	$0.01 r_{\max}$
Factor voxel size map merge α	0.5
Factor voxel size registration β	1.5
ICP convergence criterion γ	10^{-4}

of point-to-point residuals

$$\Delta T_{\text{est},j} = \underset{T}{\operatorname{argmin}} \sum_{(s,q) \in \mathcal{C}(\tau_t)} \rho(\|Ts - q\|_2), \quad (12)$$

where $\mathcal{C}(\tau_t)$ is the set of nearest neighbor correspondences with a distance smaller than τ_t and ρ is the Geman-McClure robust kernel, i.e., an M-estimator with a strong outlier rejection property, given by

$$\rho(e) = \frac{e^2/2}{\underbrace{\sigma_t/3 + e^2}_{\kappa_t}}, \quad (13)$$

where the scale parameter κ_t of the kernel is adapted online using σ_t . Lastly, we update the points s_i , i.e.,

$$\{s_i \leftarrow \Delta T_{\text{est},j} s_i \mid s_i \in \mathcal{S}\}, \quad (14)$$

and repeat the process until the convergence criterion is met.

As a result of this process, we obtain the transformation $T_t = \Delta T_{\text{icp},t} T_{t-1} T_{\text{pred},t}$, where $\Delta T_{\text{icp},t} = \prod_j \Delta T_{\text{est},j}$. While we apply the prediction model $T_{\text{pred},t}$ (i.e., the constant velocity prediction) to the local coordinate frame of the scan, we perform the ICP correction $\Delta T_{\text{icp},t}$ in the global reference frame of the robot. This is done for efficiency reasons as it allows us to transform the source points \mathcal{S} only once per ICP iteration. With this, the local pose deviation ΔT_t at time t used in the (5) can be expressed as

$$\Delta T_t = (T_{t-1} T_{\text{pred},t})^{-1} \Delta T_{\text{icp},t} T_{t-1} T_{\text{pred},t}. \quad (15)$$

A standard termination criterion for the ICP algorithm is to control the number of iterations. Additionally, most approaches also have a further criterion based on the minimum change in the solution. Conversely, we found that controlling the number of iterations does not allow the algorithm to always find a good solution. Thus, we only employ the termination criterion based on the applied correction being smaller than γ , without imposing a maximum number of iterations.

Finally, the ICP correction is applied to the point cloud $\mathcal{P}_{\text{merge}}^*$, and the points are integrated into the local map.

F. Parameters

Our implementation depends on a small set of seven parameters. All are shown in Table I. We use the same parameters for all experiments. Most other approaches use a substantially larger set of parameters: MULLS [21] has 107 parameters, SuMa [1] has 49 parameters, and CT-ICP [10] has 30 parameters in their respective configuration files. In contrast, our approach only has two parameters for the correspondence search, four for the map representation and scan subsampling, and one for the ICP

termination. Note that the maximum range of a scanner is a value that depends on the specific sensor in use and, as such, we do not consider it a system parameter. However, for some scenarios, the value of r_{\max} might also be adapted to the specific environment in which the system is operating, e.g., not considering far away measurements that are usually less accurate.

IV. EXPERIMENTAL EVALUATION

This work provides a simple yet effective LiDAR odometry pipeline that comes with a small set of parameters. We present our experiments to show the capabilities of our method. The results of our experiments support our key claims, namely that our approach (i) is on par with more complex state-of-the-art odometry systems, (ii) can accurately compute the robot's odometry in a large variety of environments and motion profiles with the same system configuration, and (iii) provides an effective solution to motion distortion without relying on IMUs or wheel odometers.

A. Experimental Setup

We use numerous datasets and common evaluation methods. We start with the KITTI odometry dataset [12] to evaluate our system against state-of-the-art approaches to LiDAR odometry. To investigate how we perform in other autonomous driving datasets employing a different sensor, we evaluate our approach on the MulRan dataset [15]. Additionally, we show that our approach can be used in different scenarios, such as the one present in the NCLT dataset [5], a segway dataset, and the Newer College dataset [22] recorded using a handheld device. We also analyze our method's different components, such as the motion-compensation scheme and the adaptive threshold.

Please note that due to space limitations we omit to show the results of the trajectories and a detailed runtime evaluation in this manuscript but refer the reader to the official project page where all the plots and per-sequence evaluation on the runtime performances are available.¹

B. Performance on the KITTI-Odometry Benchmark

This experiment evaluates the performance of different odometry pipelines on the popular KITTI benchmark dataset. Since most systems do not do motion compensation, we use the already compensated KITTI scans for a fair comparison and disable the motion compensation for our approach and CT-ICP [10] in this first analysis (the performance of the motion compensation module will be studied later in Section IV-D1). Table II exhibits how our system challenges most state-of-the-art systems, which are typically more sophisticated than our point-to-point ICP. Based on the official KITTI Benchmark, we rank second among the open-source approaches (behind CT-ICP [10]) and ninth among all submissions. This indicates that our comparably simple system still performs better than all the publicly available systems out there, except CT-ICP [10]. Note that CT-ICP is a complete SLAM system, and it uses loop closures to correct for the accumulated drift of the odometry estimation. We in contrast obtain our results using only open-loop registration without any loop closing.

¹[Online]. Available: <https://www.github.com/PRBonn/kiss-icp/tree/main/evaluation>

TABLE II
KITTI BENCHMARK RESULTS WITH MOTION COMPENSATED DATA

	Method	Seq. 00-10	Seq. 11-21
SLAM	SuMa++ [1]	0.70	1.06
	MULLS [21]	0.52	-
	CT-ICP [10]	0.53	0.59
Odometry	IMLS-SLAM [11]	0.55	0.69
	MULLS [21]	0.55	0.65
	F-LOAM [33]	0.84	1.87
	SuMa [1]	0.80	1.39
	Ours	0.50	0.61

We report the average relative translational error in % [13], we compare across slam methods employing pose-graph optimization for improved results (top) and odometry methods (bottom). we omit the relative rotational error, but these results are available at

https://www.cvlibs.net/datasets/kitti/eval_odometry.php

TABLE III
QUANTITATIVE RESULTS ON THE MULRAN DATASET [15]

Sequence	Method	Avg. tra.	Avg. rot.	ATE tra.	ATE rot.
KAIST	MULLS [21]	2.94	0.86	37.24	0.11
	SuMa [1]	5.59	1.73	43.61	0.14
	F-LOAM [33]	3.43	0.99	46.17	0.15
	Ours	2.28	0.68	17.40	0.06
DCC	MULLS [21]	2.96	0.98	38.35	0.12
	SuMa [1]	5.20	1.71	36.22	0.11
	F-LOAM [33]	3.83	1.14	42.70	0.13
	Ours	2.34	0.64	15.16	0.05
Riverside	MULLS [21]	5.42	2.21	91.16	0.16
	SuMa [1]	13.86	2.13	227.24	0.38
	F-LOAM [33]	5.47	1.18	138.09	0.22
	Ours	2.89	0.64	49.02	0.08
Sejong*	MULLS [21]	5.93	0.84	2151.00	0.49
	F-LOAM [33]	7.87	1.20	3448.97	0.82
	Ours	4.69	0.70	1369.54	0.33

we report the relative translational error and the relative rotational error using the kitti [13] metrics. additionally, we show the absolute trajectory error for translation in m and for rotation in rad.

C. Comparison to State-of-The-Art Systems on Other Datasets

We proceed to analyze the performance of our system on different datasets, scenarios, and types of robots. For that, we use the MulRan dataset [15], a handheld device [22], and a segway dataset [5]. Odometry pipelines typically deal with those challenging scenarios but employ IMUs [27] or a different system configuration [10]. Our system performs on par with state-of-the-art systems using the same parameter values for all experiments and datasets. For this experiment, we compare against state-of-the-art odometry systems, namely MULLS [21], SuMa [1], F-LOAM [33], and CT-ICP [10]. Note that we do not provide an evaluation of CT-ICP for the MulRan dataset since CT-ICP does not provide support for this dataset.

For the MulRan dataset [15], we test the systems under evaluation on all available public sequences. Since the dataset provides three similar runs for each sequence, we report the average number of each sequence in Table III. Our method outperforms all state-of-the-art approaches by a large margin in both relative and absolute error.

We use both available sequences to evaluate the Newer College dataset and achieve similar results on the short experiment

TABLE IV
QUANTITATIVE RESULTS FOR NEWER COLLEGE AND NCLT

Method	NCD 01-short	NCD 02-long	NCLT 2012-01-8
MULLS [21]	0.82	1.23	-
F-LOAM [33]	2.02	fails	-
CT-ICP [10]	0.48	0.58	1.17
Ours	0.51	0.96	1.27

We report the relative translational error in % [13].

compared to CT-ICP. For the long experiment, the performance gap can be explained by the additional loop closing module of CT-ICP, which is a complete SLAM system. For the NCLT dataset experiment, we use the sequence evaluated on the original work of CT-ICP. We could not reproduce the results reported in CT-ICP [10] and therefore report the results given in the original paper [10] in Table IV. We achieve similar results than CT-ICP. However, we observed errors in the GPS ground-truth poses and missing frames. Therefore, the numbers on NCLT should be taken with a grain of salt and rather provide an estimate of how the systems perform. We discourage using NCLT to evaluate odometry systems: misalignments in the ground truth poses, missing frames, and inconsistencies in the data make the evaluation of odometry systems on such a dataset not a good evaluation tool from our perspective. However, we provide the results for completeness.

We show qualitative results in Fig. 1 generated using our KISS-ICP poses. Using a single system configuration, we can produce consistent maps on different sensor setups (Velodyne/Ouster vs. Livox) and different motion profiles (car, drone, segway, handheld) with the same parameters.

D. Ablation Studies

To understand how each component of our system impacts the odometry performance, we conduct ablation studies on the different components of our approach, namely, the motion compensation scheme and the adaptive threshold. To carry out these studies, we use the KITTI odometry dataset [12] as it is probably the best-known one.

1) *Motion Compensation*: To assess the impact of our motion compensation scheme, we utilize the raw LiDAR point clouds without any compensation applied. Note that the KITTI odometry benchmark point cloud data [12] is already compensated and, therefore, cannot be used for this study. Thus, we use the KITTI raw dataset [13]. We present the results in a familiar fashion, selecting only the sequences that correspond to the ones on the motion-compensated datasets [12]. As we can see in Table V, our motion compensation scheme can produce state-of-the-art results and is on par with substantially more sophisticated and thus complex compensation techniques such as the one introduced by CT-ICP [10]. Additionally, we study how our system performs without applying motion compensation, as shown in Table V. We also evaluate the performance of our constant velocity model for motion compensation. To assess this, we compare the same compensation strategy but replace the velocity estimation with sensor data taken by the IMU. As seen in the results, our velocity estimation is on par or even slightly better with the IMU.

Besides the fact that CT-ICP's elastic formulation yields good results, our much simpler approach produces even better results. This result shows that the constant velocity model

TABLE V
RESULTS OF EVALUATING DIFFERENT STATE-OF-THE-ART SYSTEMS ON KITTI-RAW DATASET (WITHOUT MOTION COMPENSATION)

Method	Avg. tra	Avg. rot	Avg. freq.
MULLS [21]	1.41	-	12 Hz
IMLS-SLAM [11]	0.71	-	1 Hz
CT-ICP [10]	0.55	-	15 Hz
Ours without deskewing	0.91	0.27	51 Hz
Ours + Deskewing (IMU)	0.51	0.19	38 Hz
Ours + Deskewing (CV)	0.49	0.16	38 Hz

We report the relative translational error and the relative rotational error using the kitti [13] metrics. additionally, we report the runtime operation of the systems being in consideration for this experiment.

TABLE VI
COMPARISON OF DIFFERENT FIXED THRESHOLDS VS. OUR PROPOSED ADAPTIVE THRESHOLD ON THE KITTI DATASET

Dataset	Data-Association Threshold τ				
	0.3 m	0.5 m	1.0 m	2.0 m	Ours
KITTI Seq. 00	0.54	0.51	0.53	0.55	0.51
KITTI Seq. 04	0.39	0.41	0.37	0.39	0.36
KITTI Avg. Seq. 00-10	0.53	0.51	0.51	0.53	0.50

We report the relative translational error in % [13].

employed in our approach for compensating motion distortion is sufficient to cope with the slight reduction in performance when no compensation is applied. Consequently, we believe that more sophisticated techniques are unnecessary for *most* robotic odometry estimation.

2) *Adaptive Data-Association Threshold*: We finally evaluate how the adaptive threshold τ_t impacts the performance of our system by comparing it to a different set of fixed thresholds commonly used in open-source systems. To conduct this experiment, we identify the two KITTI sequences with the largest (00) and the smallest (04) average acceleration indicating different motion profiles. As we can see in Table VI, the best fixed threshold for sequence 00 is 0.5 m and 1.0 m for sequence 04. This means that a fixed threshold has to be tuned depending on the motion profile and thus to the dataset to achieve top performance. In contrast, our adaptive algorithm exploits the motion profile to estimate the threshold online, which results in on-par or better performance without the need to find a new fixed threshold for each sequence. Finally, our proposed adaptive threshold strategy achieves the best average result on the KITTI training sequences. Please note that all the experiments from this ablation study use the robust kernel. For space reasons, we omitted the results of the evaluation of our system when no kernel is employed and only report the results averaged over the sequences. Not using the kernel produces 0.67% for the translational error and 0.25% for the rotational error [13].

V. CONCLUSION

This paper presents a simple yet highly effective approach to LiDAR odometry and shows that point-to-point ICP works very well – when used properly. Our approach operates solely on point clouds and does not require an IMU, even when dealing with high-frequency driving profiles. Our approach exploits the classical point-to-point ICP to build a generic odometry system that can be employed in different challenging environments, such as highway runs, handheld devices, segways, and drones.

Moreover, the system can be used with different range-sensing technologies and scanning patterns. We only assume that point clouds are generated sequentially as the robot moves through the environment. We implemented and evaluated our approach on different datasets, provided comparisons to other existing techniques, supported all claims made in this paper, and released our code. The experiments suggest that our approach is on par with substantially more sophisticated state-of-the-art LiDAR odometry systems but relies only on a few parameters, and performs well on various datasets under different conditions with the same parameter set. Finally, our system operates faster than the sensor frame rate in all presented datasets. We believe this work will be a new baseline for future sensor odometry systems and a solid, high-performance starting point for future approaches. Our open-source code is robust and simple, easy to extend, and performs well, pushing the state-of-the-art LiDAR odometry to its limits and challenging most sophisticated systems.

ACKNOWLEDGMENT

The authors would like to thank Pierre Dellenbach for making the CT-ICP code available, which inspired our implementation. The authors would also like to thank Yue Pan for helping with the evaluation of MULLS for this paper and Igor Bogoslavskyi for his feedback.

REFERENCES

- [1] J. Behley and C. Stachniss, "Efficient surfel-based SLAM using 3D laser range data in urban environments," in *Proc. Robot.: Sci. Syst.*, 2018, Art. no. 59.
- [2] J. Bentley, "Multidimensional binary search trees used for associative searching," *Commun. ACM*, vol. 18, no. 9, pp. 509–517, 1975.
- [3] P. Besl and N. McKay, "A method for registration of 3D shapes," *IEEE Trans. Pattern Analysis Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992.
- [4] J. L. Blanco-Claraco, "Mobile robot programming toolkit (MRPT)," 2014. [Online]. Available: <http://www.mrpt.org/>
- [5] N. Carlevaris-Bianco, A. Ushani, and R. Eustice, "University of Michigan North Campus long-term vision and LiDAR dataset," *Int. J. Robot. Res.*, vol. 35, no. 9, pp. 1023–1035, 2016.
- [6] N. Chebrolu, T. L  be, O. Vysotska, J. Behley, and C. Stachniss, "Adaptive robust kernels for non-linear least squares problems," *IEEE Robot. Automat. Lett.*, vol. 6, no. 2, pp. 2240–2247, Apr. 2021.
- [7] X. Chen et al., "Moving object segmentation in 3D LiDAR data: A learning-based approach exploiting sequential data," *IEEE Robot. Automat. Lett.*, vol. 6, no. 4, pp. 6529–6536, Oct. 2021.
- [8] X. Chen, A. Milioto, E. Palazzolo, P. Gigu  re, J. Behley, and C. Stachniss, "SuMa: Efficient LiDAR-based semantic SLAM," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2019, pp. 4530–4537.
- [9] Y. Chen and G. Medioni, "Object modeling by registration of multiple range images," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 1991, pp. 2724–2729.
- [10] P. Dellenbach, J. Deschaud, B. Jacquet, and F. Goulette, "CT-ICP real-time elastic LiDAR odometry with loop closure," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2022, pp. 5580–5586.
- [11] J. Deschaud, "IMLS-SLAM: Scan-to-model matching based on 3D data," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 2480–2485.
- [12] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the KITTI vision benchmark suite," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 3354–3361.
- [13] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1231–1237, 2013.
- [14] T. Guadagnino, X. Chen, M. Sodano, J. Behley, G. Grisetti, and C. Stachniss, "Fast sparse LiDAR odometry using self-supervised feature selection on intensity images," *IEEE Robot. Automat. Lett.*, vol. 7, no. 3, pp. 7597–7604, Jul. 2022.
- [15] J. Jeong, Y. Cho, Y. Shin, H. Roh, and A. Kim, "Complex urban LiDAR data set," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 6344–6351.
- [16] R. Kolluri, "Provably good moving least squares," *ACM Trans. Algorithms*, vol. 4, no. 2, pp. 1–25, 2008.
- [17] J. Lin, F. Zhang, and A. Loam_livox, "Robust LiDAR odometry and mapping LOAM package for livox LiDAR," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2019, pp. 3126–3131.
- [18] K. Museth et al., "OpenVDB: An open-source data structure and toolkit for high-resolution volumes," in *Proc. ACM SIGGRAPH Courses*, 2013, Art. no. 19.
- [19] R. A. Newcombe et al., "KinectFusion: Real-time dense surface mapping and tracking," in *Proc. IEEE Int. Symp. Mixed Augmented Reality*, 2011, pp. 127–136.
- [20] M. Nie  ner, M. Zollh  fer, S. Izadi, and M. Stamminger, "Real-time 3D reconstruction at scale using voxel hashing," *ACM Trans. Graph.*, vol. 32, 2013, Art. no. 169.
- [21] Y. Pan, P. Xiao, Y. He, Z. Shao, and Z. Li, "MULLS: Versatile LiDAR SLAM via multi-metric linear least square," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2021, pp. 11633–11640.
- [22] M. Ramezani, Y. Wang, M. Camurri, D. Wisth, M. Mattamala, and M. Fallon, "The newer college dataset: Handheld lidar, inertial and vision with ground truth," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2020, pp. 4353–4360.
- [23] V. Reijgwart, A. Millane, H. Oleynikova, R. Siegwart, C. Cadena, and J. Nieto, "Voxgraph: Globally consistent, volumetric mapping using signed distance function submaps," *IEEE Robot. Automat. Lett.*, vol. 5, no. 1, pp. 227–234, Jan. 2019.
- [24] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in *Proc. IEEE Int. Conf. 3-D Digit. Imag. Model.*, 2001, pp. 145–152.
- [25] R. B. Rusu and S. Cousins, "3D is here: Point Cloud Library (PCL)," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2011, pp. 1–4.
- [26] J. Serafin and G. Grisetti, "NICP: Dense normal based point cloud registration," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2015, pp. 742–749.
- [27] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and D. Rus, "LIO-SAM tightly-coupled LiDAR inertial odometry via smoothing and mapping," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2020, pp. 5135–5142.
- [28] T. Shan and B. Englot, "LeGO-LOAM: Lightweight and ground-optimized LiDAR odometry and mapping on variable terrain," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2018, pp. 4758–4765.
- [29] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Cambridge, MA, USA: MIT Press, 2005.
- [30] E. Vespa, N. Nikolov, M. Grimm, L. Nardi, P. Kelly, and S. Leutenegger, "Efficient octree-based volumetric SLAM supporting signed-distance and occupancy mapping," *IEEE Robot. Automat. Lett.*, vol. 3, no. 2, pp. 1144–1151, Apr. 2018.
- [31] I. Vizzo, X. Chen, N. Chebrolu, J. Behley, and C. Stachniss, "Poisson surface reconstruction for LiDAR odometry and mapping," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2021, pp. 5624–5630.
- [32] I. Vizzo, T. Guadagnino, J. Behley, and C. Stachniss, "VDBFusion: Flexible and efficient TSDF integration of range sensor data," *Sensors*, vol. 22, no. 3, 2022, Art. no. 1296.
- [33] H. Wang, C. Wang, C. Chen, and L. Xie, "F-LOAM: Fast LiDAR odometry and mapping," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2021, pp. 4390–4396.
- [34] M. Zeng, F. Zhao, J. Zheng, and X. Liu, "Octree-based fusion for realtime 3 d reconstruction," *Graphical Models*, vol. 75, no. 3, pp. 126–136, 2013.
- [35] J. Zhang and S. Singh, "LOAM: LiDAR odometry and mapping in real-time," in *Proc. Robotics: Sci. Syst.*, 2014, pp. 181–185.
- [36] Q. Zhou, J. Park, and V. Koltun, "Open3D: A modern library for 3D data processing," 2018, *arXiv:1801.09847*.