Eötvös Loránd University

Faculty of Informatics

Erasmus Mundus Joint Master in

Intelligent Field Robotic Systems

LiDAR Odometry and Mapping Beyond RTK Accuracy

Author:

Liviu-Daniel Florescu

MSc Intelligent Field Robotic Systems (IFRoS)

Internal supervisor:

Iván Eichhardt

Assistant Professor

External supervisor:

Maximilian Fenkart

CTO, Sodex Innovations GmbH

Budapest, 2025

Title of the thesis

LiDAR Odometry and Mapping Beyond RTK Accuracy

Topic of the thesis

Out of the five basic senses that humans use to experience the world, vision accounts for 80% of the information input that our brains operate with [1]. Naturally, robotics research aims to replicate this and develop systems that can not only collect high-quality visual data but also create rich artificial representations of the world, enabling autonomous systems to confidently reason about their environment.

This work will focus on the use of 3D Light Detection and Ranging (LiDAR) sensors in outdoor settings. When mounted on an arbitrary mobile base and moved around a target environment, the sensor collects information about the geometry of the scene, which can be merged in order to create a general 3D model. However, this process depends on accurate displacement measurements that are not trivial to obtain. An existing solution relies on Global Navigation Satellite System (GNSS) localization, corrected using Real-Time Kinematic positioning (RTK), and orientation from an Inertial Measurement Unit (IMU). Together, these create an Inertial Navigation System (INS) whose output can be interpreted as the 3D transformation between sensor poses at discrete time steps. Even though this represents the state-of-the-art technology for outdoor localization, with centimeter-level position error, its accuracy is unsatisfactory when it comes to high-quality pointcloud registration. Additionally, this is not feasible for all outdoor scenes, because GNSS accuracy varies heavily depending on surroundings and signal strength.

We will investigate methods that address the limitations of the INS-based registration, by using the visual information in the scene, such that the system is less reliant on a sensor with fluctuating uncertainty. Previous research in this area [2] indicates that visual cues alone should be enough to achieve reliable displacement estimation, enabling the computation of odometry from LiDAR data, as well as creating a 3D map of the explored environment. A comparison between LiDAR odometry and GNSS localization is also within the scope of this work. The research questions that we aim to answer are the following:

- What metrics exist for measuring the accuracy of pointcloud registration?
- Can methods that use only visual information achieve higher quality pointcloud registration (3D mapping) than RTK-based merging?
- To what extent is LiDAR-based odometry an alternative to GNSS localization?

The work will be carried out in collaboration with Sodex Innovations GmbH, who provide the sensor rig (LiDAR, RGB cameras, GNSS + RTK + IMU) as well as several existing datasets consisting of sensor data collected while exploring outdoor rural environments. The project will span approximately 15 weeks, and is tentatively structured as shown below, subject to change brought by results at different stages.

Stage	Estimated Duration	Description
Introduction	2 weeks	Familiarization with the existing sensors, datasets and related processes
Background investigation	3 weeks	Exploring existing solutions, initial experimentation on available data
Design and Development	4 weeks	Iterative implementation and testing of new approaches
Analysis	4 weeks	Evaluation and comparison between the solutions developed at different iterations
Final write-up	2 weeks	Reporting methods, implementation and results

Keywords
LiDAR odometry, LiDAR mapping, pointcloud registration, GNSS, RTK

References

[1] Man, Dariusz & Olchawa, Ryszard. (2018). The Possibilities of Using BCI Technology in Biomedical Engineering. 10.1007/978-3-319-75025-5_4.
[2] Lee, Dongjae & Jung, Minwoo & Yang, Wooseong & Kim, Ayoung. (2024). LiDAR odometry survey: recent advancements and remaining challenges. Intelligent Service Robotics. 17. 1-24. 10.1007/s11370-024-00515-8.

Contents

Chapter 1

Introduction

Like many scientific terms that we encounter in our daily activities, *robotics* describes a broad collection of technologies and stands at the intersection of key research directions. Motivated by practicality, disciplines that appear completely unrelated find themselves building upon advancements in other fields, diluting boundaries and joining forces to enable otherwise-impossible advancements.

What constitutes a *robot*, then? In its inceptive use by Karel Čapek [1] in 1921, the word stemmed from the Slavic *robota*, meaning “servitude” or “forced work”, and referred to a human-like mechanical system working on factory assembly lines. This concept, however, dates from earlier centuries. Around 1495, Leonardo Da Vinci envisioned a mechanical knight controlled by a series of pulleys, that was able to perform simple movements [2]. Testifying to the industrial revolution and the computational breakthroughs of the last few decades, modern-day humanoid robots can perform acrobatics [3], interact with humans in constrained scenarios [4] [5] [6] and even replicate human facial expressions [7]. Still, a device ought not necessarily appear human-like in order for it to be labeled as a robot. Autonomous vacuum cleaners, space rovers or crop-monitoring drones fall under the same category. At this stage, a complete taxonomy would have to address dozens of physical (size, shape, mobility, locomotion system, etc.) and non-physical (autonomy, perception abilities, use-case, etc.) characteristics, and none of these would independently convey the meaning that we intuitively associate to the notion of *robot*. Without assessing whether an exhaustive, generic definition is even possible, we can synthesize the above by affirming that a *robot* is an artificial system that performs one or more tasks and is able to evaluate its state or gather information from its environment.

1.1 Robotic perception

This formulation highlights two essential aspects of a robotic agent (Fig. 1.1): actuation, seen as some form of dynamic physical ability, allowing the agent to enact the desired behavior, and perception, the ability to observe changes in the environment.

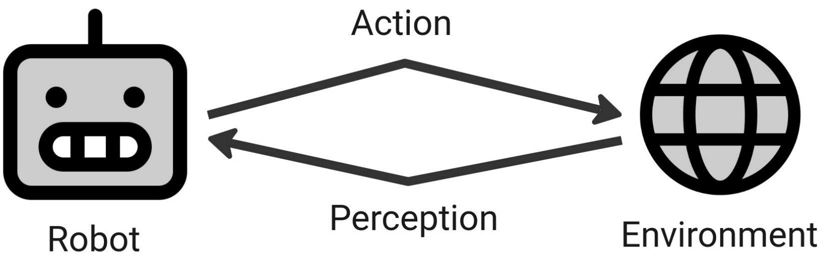

Figure 1.1: A simplified interpretation of the robotics paradigm. The interaction between an agent and its environment can be seen as a two-way flow: the agent alters the environment through its actions, and uses perception to observe it.

Sensing modalities have largely different contributions to the perception mechanism. To this extent, the phrase *visual dominance* was introduced by F. Colavita [8], whose study demonstrated that humans focus more on the visual component when presented with an audiovisual stimulus, and following research has strongly confirmed this tendency [9] [10]. Unsurprisingly, a similar pattern is emerging in the case of robots, thanks to the reduced cost, familiarity and wide availability of cameras. In many situations, visual stimuli provide most of the necessary information, and this has motivated the development of various image processing algorithms.

Technological innovation in the last century has led to the situation in which artificial sensors surpass humans in both the range of signals that are perceived, as well as the accuracy of the measurements. A relevant example is the class of Light Detection and Ranging (LiDAR) sensors which retrieve three-dimensional information about the environment at a very high frequency and with rather negligible measurement errors, in the form of *point clouds*. Among many applications, this type of sensors can be used to construct virtual representations of a specific environment, enabling engineers to experiment with a realistic model, evaluate construction progress or validate a finished project.

Figure 1.2: The SDX-Compact manufactured by Sodex Innovations GmbH. The set of sensors consists of a 3D LiDAR scanner, three RGB cameras and a high-accuracy positioning system. Image source: Fieldwork

1.2 Problem definition

The current work addresses a common and well-known problem in the area of Field Robotics, namely Simultaneous Localization and Mapping (SLAM), and combines the practicality of an industrial solution with a perception-based approach.

SDX-Compact (Fig. 1.2) is the main product of Sodex Innovations GmbH, consisting of multiple sensors that collect spatial and visual data. This module can be easily mounted on an arbitrary vehicle in order to expand its perception capabilities and convert it into a surveying device. While the vehicle is moving, the LiDAR sensor captures 3D scans of the surroundings, as well as related metadata (timestamps, localization information, signal intensity etc.). Because the rig includes a high-accuracy Inertial Navigation System (INS), the localization and orientation data can be used to join the collected point clouds and create a global 3D map of the traversed space.

Nonetheless, this suffers from two main limitations:

- Reliance on unstable signal: internally, the INS depends on information from a Global Navigation Satellite System (GNSS) receiver, which is limited to outdoor spaces and whose availability varies depending on weather conditions and surroundings (e.g. thick vegetation, tall buildings, bridges).
- Unsatisfactory accuracy: when merging point cloud data, positioning or localization errors introduce inconsistencies in the final 3D model, which hinders precise planning and construction.

Our work aims to address these limitations by introducing a component that utilizes the information collected by the LiDAR sensor in conjunction with the existing data.

Three research questions have been formulated to guide this process:

1. What metrics exist for measuring the accuracy of point cloud registration? In this context, registration refers to placing a pair of related point sets in a common reference frame.
2. Can methods that use only visual information achieve higher quality point cloud registration (3D mapping) than merging based on Real-Time Kinematics (RTK)? Usually, GNSS systems provide meter-level accuracy. In the current scenario, however, the system is corrected using Real-Time Kinematics, such that the expected error is at centimeter-level.
3. To what extent is LiDAR-based odometry an alternative to GNSS localization? Previous research indicates that the spatial information present in 3D point clouds could be used to compute the relative displacement between consecutive scans, resulting in the ability to estimate odometry (an essential component of robotic localization) without dedicated sensors such as wheel encoders or accelerometers.

The main contribution of the project consists of developing an original framework for localization and mapping based on data collected with an industrial sensor rig. The results are more generic than if a particular physical robotic system were involved, and thus are relevant for virtually any robotic application with a similar setup.

The following chapters of this document will cover related research directions and efforts that our work builds upon (Chapter 2), a detailed description of the components and algorithms involved in developing the project (Chapter 3), an evaluation of the method based on its results (Chapter 4), as well as a series of conclusions that were drawn from the overall process (Chapter 5).

Chapter 2

Background and Literature Review

The goal of this chapter is to present the research context in which our project was conducted, by looking at common methods for each of the main components and highlighting those that inspired the current approach.

2.1 Simultaneous Localization and Mapping (SLAM)

SLAM constitutes a key research area in robotics, because it is a foundational building block for autonomous operation. This problem occurs when the robot does not have prior access to a map of the environment, so it must construct one while keeping track of its current position (*online* SLAM). If the goal is to optimize the entire sequence of poses along the robot path, this is known as the *full* SLAM problem. [11]

Usually, the problem is discretized along the time dimension, such that at time t we aim to compute the posterior function $p(x_t, M|z_{1:t}, u_{1:t})$, where x_t is the current state, M is the map, $z_{1:t}$ represents the set of measurements collected so far, and $u_{1:t}$ the control sequence. As each of these variables can have different concrete representations, depending on the task and the available sensors, a broad range of approaches have been proposed.

Aulinas et al. [12] note that the earliest methods rely on probabilistic models derived from the recursive Bayes rule, as such formulations can provide intuitive representations of the various noisy components involved in a robotic system. When employing the Kalman Filter (KF) [13] or its variations, the robot state, measurements and control inputs are modeled as multi-dimensional Gaussians, whose covariances describe the associated uncertainty. The algorithm alternates between *predictions*, when the state is modified based on u_t, and *updates*, when the prediction is evaluated against the latest observation z_t, and the state hypothesis is updated accordingly. The Extended Kalman Filter (EKF) maintains this structure but can accommodate non-linearities in the displacement or measurement functions by using local linear approximations. Such methods have been successfully applied for indoor [14], aerial [15], and underwater [16] robots.

Particle Filter (PF), introduced by Del Moral [17], is a probabilistic approach that relies on the Monte Carlo method. Instead of an analytical form, the uncertainty is accounted for by considering a large set of samples (representing potential states) and weighting them based on measurement likelihood. This has a higher computational cost than the standard KF methods, but is not affected by any linearization limitations. Nie et al. [18] implemented a LiDAR SLAM algorithm based on PF localization, albeit for 2D mapping.

Another large group of methods is represented by *Visual SLAM*, when cameras are the main (and sometimes only) sensor used. As early as 1980, Moravec [19] developed a robot capable of estimating its motion by matching features in images captured at discrete poses, in a stop-and-go fashion. In 2007, SLAM was being performed using a single handheld camera [20] [21], triggering a separation from the popular, offline, Structure from Motion (SfM) techniques. The solution is even more robust when a stereo system is available [22], as this helps avoid the geometrical limitations of monocular vision.

2.2 Point cloud Registration

Before discussing LiDAR-based approaches in more detail, let us review the task of point cloud registration. Given two sets of points P, Q in arbitrary reference frames, representing (at least partially) the same scene, the goal is to find the transformation $T \in SE(3)$ (translation and rotation) that best aligns the points. This can be expressed as an optimization problem:

$$\underset{T}{\operatorname{argmin}}\, J(TP, Q) \tag{2.1}$$

where J is a custom cost function. The particular case where known point correspondences are provided has a closed-form solution [23], but the Iterative Closest Point (ICP) algorithm [24] removes this constraint by alternating between correspondence generation — each point in P is paired with its nearest neighbor from Q — and minimizing the point-pair distances. This approach is by far the most widespread, thanks to its simplicity, computational efficiency (e.g. using a k-d tree for closest-point search) and many available variations [25] [26].

Chen and Medioni [27] proposed a "point-to-plane" minimization, while Segal et al. [28] formulated a generalized cost function by considering the probabilistic model underlying the point sets, and implemented a "plane-to-plane" minimization. Other variants modify the point selection strategy [29] [30], apply a weight to each pair [31], prune the set of the correspondences [32] [33] or employ a different minimization algorithm [34].

A different class of solutions operates by discretizing the space into voxels and estimating a normal distribution using the points in each voxel. Introduced in 2003, the Normal Distributions Transform (NDT) [35] algorithm is the main competitor to ICP, enabling registration to be performed without the need for point-to-point correspondences. Magnusson et al. [36] extended this work to 3D scans, and performed a thorough comparison between this and ICP [37] for the challenging task of registering point clouds captured in underground mines, where few usable features are present. More recently, the method has been applied for outdoor mapping [38], and as the basis for a Lidar Odometry and Mapping solution [39].

Neural Networks can also be used for point cloud registration, by creating correspondences based on features extracted by a descriptor architecture [40] [41]. This is particularly useful when no initial estimate of the transformation between the two point sets is available — a scenario that the other approaches cannot directly handle.

2.3 LiDAR-based Odometry and Mapping

Despite their high price range, LiDAR sensors have gained significant popularity in the last decade, becoming the go-to solution for autonomous mobility. Unlike cameras, LiDARs are not dependent on ambient lighting, so they can operate in total darkness, and they provide high-frequency, reliable 3D information without the need for additional processing. Basic applications include obstacle detection [42] [43], but the amount of information they provide makes them excellent for localization and mapping purposes.

A key reference is LiDAR Odometry and Mapping (LOAM), the work of Zhang and Singh [44], who used a motorized, 2-axis LiDAR scanner to generate 3D pointclouds. To register consecutive scans, they perform feature extraction (sharp edges or planar patches), identify correspondences to the previous scan, and find the optimal transformation using the Levenberg-Marquardt method. To account for displacement during the beam sweep, points are reprojected by linear interpolation, assuming constant velocity — this is known as *motion/distortion compensation*. Mapping takes place in parallel, at a slower rate, and the current scan is registered to the map created so far. At that time, this implementation achieved the best results [1] on the KITTI Odometry Benchmark [45]. V-LOAM [46] improved this solution by introducing a visual odometry component based on a fish-eye camera.

A few years later, LeGO-LOAM [47] extended the feature extraction process by performing segmentation on a range image generated from the 3D point cloud. Clusters are filtered by size, in order to discard features belonging to potentially noisy points. Another modification is that the parameters of the transformation between consecutive scans are optimized separately: $t_z, \theta_{roll}, \theta_{pitch}$ are computed using planar feature correspondences, then fixed during the optimization of t_x, t_y, θ_{yaw}, which only uses edge features. The method is compared to LOAM and achieves higher accuracy in outdoor scenarios, while being an order of magnitude faster.

F-LOAM [48] proposes a two-step motion compensation process. For odometry computation, the constant velocity model is used, but the points are re-corrected using the optimized pose before being registered to the map. Correspondences are weighted based on the "local smoothness" of the feature, and the non-linear optimization is solved using the Gauss-Newton method. The map is not updated at every scan, but only when the translational change reaches a predefined threshold. This is a common keyframe selection technique inherited from Visual Odometry.

A slightly different approach introduces the use of inertial sensors, leading to LiDAR-Inertial Odometry (LIO). This aims to compensate for the lack of reliable 3D features in specific environments, as accelerometers can provide satisfactory motion estimates for short displacements. FAST-LIO [49] applies such a technique for a drone equipped with a high-frequency solid-state LiDAR, by performing tightly-coupled fusion: instead of using the IMU output to correct the scan registration, it is applied on the features extracted from the point cloud. COIN-LIO [50] introduces a photometric error component, based on the beam intensity values returned by the LiDAR. A monochrome "intensity image" is constructed and filtered such that matching can occur between consecutive frames, and these new correspondences extend the residual vector that is minimized for odometry computation. The approach achieves state-of-the-art performance on a dataset of

[1] `https://www.cvlibs.net/datasets/kitti/eval_odometry.php`

geometrically-degenerate scenes (ENWIDE [2]).

In the class of deep-learning techniques, we highlight Deep Matching LiDAR Odometry (DMLO) [51], which translates the registration problem into a supervised machine learning task. Point clouds are projected into a 2D map using cylinder encoding, with range and intensity values as channels, and a Convolutional Neural Network (CNN) architecture is trained to predict correspondences from pairs of projections. Training samples are constructed from a subset of the target dataset, and the method proves robust across different LiDAR hardware, but does not surpass LOAM on the KITTI sequences. In comparison to [52], this solution does not leave the geometric problem to the inner workings of the CNN.

The approach that our work draws most inspiration from is KISS-ICP [53], a LiDAR-only odometry framework built around point-to-point ICP. This method proposes a few modifications that cooperate towards a hardware-agnostic solution, with a small number of adjustable parameters. The first contribution is a constant velocity motion estimation, which provides the optimization step with an initial guess. The same motion estimation is used for distortion compensation. Secondly, feature extraction is replaced by two stages of voxel-based down-sampling: the first down-sampled point cloud is used to extend the map, while the even lower-resolution set is registered against the existing map to compute the pose estimate. Perhaps the key component is an adaptive distance threshold for correspondence outlier removal. This threshold is updated based on the deviation between the predicted and optimized pose, acting as a form of uncertainty estimation. Additionally, the optimization problem employs a robust kernel, whose scale parameter is related to the adaptive threshold.

[2] `https://projects.asl.ethz.ch/datasets/enwide`

Chapter 3

Methodology

Chapter 4

Results

Chapter 5

Conclusion

Bibliography

[1] Eric Roberts. *Robotics: A Brief History*. https://cs.stanford.edu/people/eroberts/courses/soco/projects/1998-99/robotics/history.html. Accessed: 2025-02-01. 1999.

[2] Mario Taddei. *Leonardo's Robots*. https://tinyurl.com/mt2a4s9x. Archived version accessed: 2025-02-01. 2008.

[3] Boston Dynamics. *Leaps, Bounds, and Backflips*. https://bostondynamics.com/blog/leaps-bounds-and-backflips/. Accessed: 2025-02-01. 2023.

[4] Subhodeep Mukherjee et al. "Humanoid robot in healthcare: A Systematic Review and Future Research Directions". In: *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*. Vol. 1. 2022, pp. 822–826. DOI: 10.1109/COM-IT-CON54601.2022.9850577.

[5] Kai-Yi Chin, Chin-Hsien Wu, and Zeng-Wei Hong. "A Humanoid Robot as a Teaching Assistant for Primary Education". In: *2011 Fifth International Conference on Genetic and Evolutionary Computing*. 2011, pp. 21–24. DOI: 10.1109/ICGEC.2011.13.

[6] Uvais Qidwai, Saad Bin Abul Kashem, and Olcay Conor. "Humanoid Robot as a Teacher's Assistant: Helping Children with Autism to Learn Social and Academic Skills". In: *Journal of Intelligent & Robotic Systems* 98.3 (2020), pp. 759–770. ISSN: 1573-0409. DOI: 10.1007/s10846-019-01075-1. URL: https://doi.org/10.1007/s10846-019-01075-1.

[7] Do Hyoung Kim et al. "Development of a Facial Expression Imitation System". In: *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2006, pp. 3107–3112. DOI: 10.1109/IROS.2006.282329.

[8] Francis B Colavita. "Human sensory dominance". In: *Perception & Psychophysics* 16.2 (1974), pp. 409–412.

[9] Fabian Hutmacher. "Why Is There So Much More Research on Vision Than on Any Other Sensory Modality?" In: *Frontiers in Psychology* 10 (2019), p. 2246. DOI: `10.3389/fpsyg.2019.02246`. URL: `https://doi.org/10.3389/fpsyg.2019.02246`.

[10] David Hecht and Miriam Reiner. "Sensory dominance in combinations of audio, visual and haptic stimuli". In: *Experimental brain research* 193 (2009), pp. 307–314.

[11] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, 2005. ISBN: 0262201623.

[12] Josep Aulinas et al. "The SLAM problem: a survey". In: *Artificial Intelligence Research and Development* (2008), pp. 363–371.

[13] Rudolph Emil Kalman. "A New Approach to Linear Filtering and Prediction Problems". In: *Transactions of the ASME–Journal of Basic Engineering* 82.Series D (1960), pp. 35–45.

[14] A.J. Davison and D.W. Murray. "Simultaneous localization and map-building using active vision". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24.7 (2002), pp. 865–880. DOI: `10.1109/TPAMI.2002.1017615`.

[15] Chunbo Luo et al. "UAV position estimation and collision avoidance using the extended Kalman filter". In: *IEEE Transactions on vehicular technology* 62.6 (2013), pp. 2749–2762.

[16] Albert Palomer, Pere Ridao, and David Ribas. "Inspection of an underwater structure using point-cloud SLAM with an AUV and a laser scanner". In: *Journal of field robotics* 36.8 (2019), pp. 1333–1344.

[17] Pierre Del Moral. "Nonlinear filtering: Interacting particle resolution". In: *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics* 325.6 (1997), pp. 653–658.

[18] Fuyu Nie et al. "LCPF: A Particle Filter Lidar SLAM System With Loop Detection and Correction". In: *IEEE Access* 8 (2020), pp. 20401–20412. DOI: `10.1109/ACCESS.2020.2968353`.

[19] Hans Peter Moravec. *Obstacle avoidance and navigation in the real world by a seeing robot rover*. Stanford University, 1980.

[20] Andrew J. Davison et al. "MonoSLAM: Real-Time Single Camera SLAM". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.6 (2007), pp. 1052–1067. DOI: `10.1109/TPAMI.2007.1049`.

[21] Georg Klein and David Murray. “Parallel tracking and mapping for small AR workspaces”. In: *2007 6th IEEE and ACM international symposium on mixed and augmented reality*. IEEE. 2007, pp. 225–234.

[22] Christopher Mei et al. “RSLAM: A system for large-scale mapping in constant-time using stereo”. In: *International journal of computer vision* 94 (2011), pp. 198–214.

[23] K. S. Arun, T. S. Huang, and S. D. Blostein. “Least-Squares Fitting of Two 3-D Point Sets”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-9.5 (1987), pp. 698–700. DOI: 10.1109/TPAMI.1987.4767965.

[24] Paul J Besl and Neil D McKay. “Method for registration of 3-D shapes”. In: *Sensor fusion IV: control paradigms and data structures*. Vol. 1611. Spie. 1992, pp. 586–606.

[25] Szymon Rusinkiewicz and Marc Levoy. “Efficient variants of the ICP algorithm”. In: *Proceedings third international conference on 3-D digital imaging and modeling*. IEEE. 2001, pp. 145–152.

[26] Xiaoshui Huang et al. *A comprehensive survey on point cloud registration*. 2021. arXiv: 2103.02690 [cs.CV]. URL: https://arxiv.org/abs/2103.02690.

[27] Y. Chen and G. Medioni. “Object modeling by registration of multiple range images”. In: *Proceedings. 1991 IEEE International Conference on Robotics and Automation*. 1991, 2724–2729 vol.3. DOI: 10.1109/ROBOT.1991.132043.

[28] Aleksandr Segal, Dirk Haehnel, and Sebastian Thrun. “Generalized-icp.” In: *Robotics: science and systems*. Vol. 2. 4. Seattle, WA. 2009, p. 435.

[29] T. Masuda, K. Sakaue, and N. Yokoya. “Registration and integration of multiple range images for 3-D model construction”. In: *Proceedings of 13th International Conference on Pattern Recognition*. Vol. 1. 1996, 879–883 vol.1. DOI: 10.1109/ICPR.1996.546150.

[30] Greg Turk and Marc Levoy. “Zippered polygon meshes from range images”. In: *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*. 1994, pp. 311–318.

[31] Guy Godin, Marc Rioux, and Rejean Baribeau. “Three-dimensional registration using range and intensity information”. In: *Videometrics III*. Vol. 2350. SPIE. 1994, pp. 279–290.

[32] Kari Pulli. “Multiview registration for large data sets”. In: *Second international conference on 3-d digital imaging and modeling (cat. no. pr00062)*. IEEE. 1999, pp. 160–168.

[33] Sofien Bouaziz, Andrea Tagliasacchi, and Mark Pauly. “Sparse iterative closest point”. In: *Computer graphics forum*. Vol. 32. 5. Wiley Online Library. 2013, pp. 113–123.

[34] Gérard Blais and Martin D. Levine. “Registering multiview range data to create 3D computer objects”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17.8 (1995), pp. 820–824.

[35] Peter Biber and Wolfgang 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. 2003, pp. 2743–2748.

[36] Martin Magnusson, Achim Lilienthal, and Tom Duckett. “Scan registration for autonomous mining vehicles using 3D-NDT”. In: *Journal of Field Robotics* 24.10 (2007), pp. 803–827.

[37] Martin Magnusson et al. “Evaluation of 3D registration reliability and speed-A comparison of ICP and NDT”. In: *2009 IEEE International Conference on Robotics and Automation*. IEEE. 2009, pp. 3907–3912.

[38] Yueqian Shen et al. “MI-NDT: Multiscale Iterative Normal Distribution Transform for Registering Large-Scale Outdoor Scans”. In: *IEEE Transactions on Geoscience and Remote Sensing* 62 (2024), pp. 1–13. DOI: 10.1109/TGRS.2024.3437162.

[39] Shoubin Chen et al. “NDT-LOAM: A real-time LiDAR odometry and mapping with weighted NDT and LFA”. In: *IEEE Sensors Journal* 22.4 (2021), pp. 3660–3671.

[40] Zan Gojcic et al. “The perfect match: 3d point cloud matching with smoothed densities”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019, pp. 5545–5554.

[41] Haowen Deng, Tolga Birdal, and Slobodan Ilic. “Ppfnet: Global context aware local features for robust 3d point matching”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 195–205.

[42] Alireza Asvadi et al. “3D Lidar-based static and moving obstacle detection in driving environments: An approach based on voxels and multi-region ground planes”. In: *Robotics and Autonomous Systems* 83 (2016), pp. 299–311.

[43] Liang Chen, Jian Yang, and Hui Kong. “Lidar-histogram for fast road and obstacle detection”. In: *2017 IEEE international conference on robotics and automation (ICRA)*. IEEE. 2017, pp. 1343–1348.

[44] Ji Zhang, Sanjiv Singh, et al. “LOAM: Lidar odometry and mapping in real-time.” In: *Robotics: Science and systems*. Vol. 2. 9. Berkeley, CA. 2014, pp. 1–9.

[45] Andreas Geiger, Philip Lenz, and Raquel Urtasun. “Are we ready for autonomous driving? the kitti vision benchmark suite”. In: *2012 IEEE conference on computer vision and pattern recognition*. IEEE. 2012, pp. 3354–3361.

[46] Ji Zhang and Sanjiv Singh. “Visual-lidar odometry and mapping: Low-drift, robust, and fast”. In: *2015 IEEE international conference on robotics and automation (ICRA)*. IEEE. 2015, pp. 2174–2181.

[47] Tixiao Shan and Brendan Englot. “LeGO-LOAM: Lightweight and Ground-Optimized Lidar Odometry and Mapping on Variable Terrain”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2018, pp. 4758–4765.

[48] Han Wang et al. “F-loam: Fast lidar odometry and mapping”. In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2021, pp. 4390–4396.

[49] Wei Xu and Fu Zhang. “FAST-LIO: A Fast, Robust LiDAR-Inertial Odometry Package by Tightly-Coupled Iterated Kalman Filter”. In: *IEEE Robotics and Automation Letters* 6.2 (2021), pp. 3317–3324. DOI: 10.1109/LRA.2021.3064227.

[50] Patrick Pfreundschuh et al. “COIN-LIO: Complementary Intensity-Augmented LiDAR Inertial Odometry”. In: *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2024, pp. 1730–1737.

[51] Zhichao Li and Naiyan Wang. “Dmlo: Deep matching lidar odometry”. In: *2020 ieee/rsj international conference on intelligent robots and systems (iros)*. IEEE. 2020, pp. 6010–6017.

[52] Qing Li et al. “Lo-net: Deep real-time lidar odometry”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019, pp. 8473–8482.

[53] Ignacio Vizzo et al. “KISS-ICP: In Defense of Point-to-Point ICP – Simple, Accurate, and Robust Registration If Done the Right Way”. In: *IEEE Robotics and Automation Letters (RA-L)* 8.2 (2023), pp. 1029–1036. DOI: `10.1109/LRA.2023.3236571`.

List of Figures

Acronyms

CNN Convolutional Neural Network.

EKF Extended Kalman Filter.

GNSS Global Navigation Satellite System.

ICP Iterative Closest Point.

INS Inertial Navigation System.

KF Kalman Filter.

LiDAR Light Detection and Ranging.

LIO LiDAR-Inertial Odometry.

LOAM LiDAR Odometry and Mapping.

NDT Normal Distributions Transform.

PF Particle Filter.

RTK Real-Time Kinematics.

SLAM Simultaneous Localization and Mapping.

Glossary

Field Robotics The area of robotics that focuses on robots operating in unstructured outdoor environments for tasks like agriculture or exploration.

k-d tree A data structure for organizing entries in a k-dimensional space using a tree structure. When k is not very large, it provides logarithmic look-up time..

keyframe An image frame or 3D scan that is used for odometry estimation. Because modern sensors usually operate at higher frequencies than needed for most algorithms, it is a reasonable decision to skip frames based on a predefined pattern or condition, without sacrificing accuracy..

odometry The process of computing relative displacement of a robot using on-board sensors.

point cloud A set of point coordinates obtained as the output of a scanning sensor.

registration The process of aligning multiple point clouds in a single coordinate system such that matching features are as close as possible.

Structure from Motion (SfM) The task of reconstructing a 3D scene and a sequence of camera poses from a set of images.

surveying The work of examining and recording the area and features of a piece of land in order to construct a map, plan, or detailed description thereof.

voxel A cuboid-shaped region in 3D space; a 3D cell..