Technische Universität München

TUM School of Computation, Information and Technology

Multi-sensor Fusion and Low-delay Ego-motion Estimation for 6D SLAM

Mojtaba Karimi

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender:		Prof. Dr.-Ing. Norbert Hanik
Prüfer der Dissertation:	1.	Prof. Dr.-Ing. Eckehard Steinbach
	2.	Prof. Dr.-Ing. Joachim Horn

Die Dissertation wurde am 15.06.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 07.11.2022 angenommen.

Abstract

In many real-world applications, 6D ego-motion estimation and mapping must be performed in real-time. Particularly in the robotics field, low-latency and robust motion estimation are essential for the control of autonomous vehicles. Dynamically generated maps are also essential for obstacle avoidance and path planning. To date, it remains a rather challenging task to fuse a variety of sensors and their substantial volumes of data in real-time. The complexity of the problem increases when the sensors are subjected to external perturbations and measurement errors. The problem is particularly hard when the ego-motion estimation and mapping are to be done in 6D, accurately, robustly, with low-latency, and in a small form factor.

In this thesis, we propose to tackle the problem by leveraging range, magnetic, and inertial sensing in a coarse-to-fine manner. The content of this work is divided into two main subsections: robust attitude and heading estimation in an indoor environment using a multi-sensor fusion approach, and low-latency 6D ego-motion estimation and mapping techniques using LiDAR-based systems. In the first part, we present a novel multi-sensor fusion approach based on skewed redundant inertial and magnetic sensors. It is developed to perform robust attitude and heading estimation with the ability to compensate for external magnetic field anomalies. We formulated a correlation-based filter model for preprocessing the sensory data and adopted an Recurrent Neural Networks (RNN) fusion model to perform robust estimation in an indoor environment. In the second part, we present a low-latency SLAM framework based on LiDAR scan slicing and concurrent matching approach. This framework uses sliced point cloud data from a rotating LiDAR in a concurrent multi-threaded matching pipeline and utilizes the attitude and heading angles to enable high update rate and low-latency 6D ego-motion estimation. A Lissajous spinning pattern is applied to overcome the sensor's limited Field of View (FOV). A two-dimensional roughness model is developed to extract the feature points for fine matching and registration of the point cloud. In addition, the pose estimator engages a temporal motion predictor that assists in finding the feature correspondences in the map for the fast convergence of the non-linear optimizer.

We have validated the proposed ego-motion estimation and mapping methods through a series of extensive experiments ranging from telepresence, hand-carried, to drone-attached setups. Throughout the experiments, different environments are explored, such as indoor labs, offices, domestic and industrial sites, as well as a wide range of mixed conditions. Results indicate that the methods are capable of carrying out high-precision, and low-latency estimation as well as robustness in terms of fast motion and environmental degradation.

Kurzfassung

In vielen realen Anwendungen müssen 6D-Ego-Motion-Schätzung und -Mapping in Echt-zeit durchgeführt werden. Gerade im Bereich der Robotik sind eine niedrige Latenz und eine robuste Bewegungsschätzung für die Steuerung autonomer Fahrzeuge unerlässlich. Dynamisch generierte Karten sind auch für die Hindernisvermeidung und Wegplanung unerlässlich. Bis heute bleibt es eine ziemlich herausfordernde Aufgabe, eine Vielzahl von Sensoren und ihre erheblichen Datenmengen in Echtzeit zu fusionieren. Die Komplexität des Problems nimmt zu, wenn die Sensoren externen Störungen und Messfehlern ausgesetzt sind. Das Problem ist besonders schwierig, wenn die Ego-Motion-Schätzung und -Zuordnung in 6d genau, robust, mit geringer Latenz und in einem kleinen Formfaktor durchgeführt werden sollen.

In dieser Arbeit schlagen wir vor, das Problem anzugehen, indem wir Entfernungs-, Magnet- und Trägheitssensorik auf grobe bis feine Weise nutzen. Der Inhalt dieser Arbeit ist in zwei Hauptunterabschnitte unterteilt: robuste Lage- und Kursschätzung in einer Innenumgebung unter Verwendung eines Multi-Sensor-Fusion-Ansatzes und 6D-Ego-Motion-Schätzung und Mapping-Techniken mit niedriger Latenzzeit unter Verwendung von Lidar-basierten Systemen. Im ersten Teil stellen wir einen neuartigen Multisensor-Fusionsansatz vor, der auf symmetrischen redundanten Inertial- und Magnetsensoren basiert. Es wurde entwickelt, um eine robuste Lage- und Steuerkursschätzung mit der Fähigkeit durchzuführen, Anomalien externer Magnetfelder zu kompensieren. Wir haben ein korrelationsbasiertes Filtermodell zur Vorverarbeitung der sensorischen Daten formuliert und ein RNN-Fusionsmodell übernommen, um eine robuste Schätzung in einer Innenumgebung durchzuführen. Im zweiten Teil stellen wir ein SLAM-Framework mit niedriger Latenz vor, das auf Lidar-Scan-Slicing und gleichzeitigem Matching-Ansatz basiert. Dieses Framework verwendet geschnittene Punktwolkendaten aus einem rotierenden Lidar in einer gleichzeitigen Multithread-Matching-Pipeline und nutzt die Lage- und Kurswinkel, um eine hohe Aktualisierungsrate und eine 6D-Ego-Motion-Schätzung mit geringer Latenz zu ermöglichen. Ein Lissajous-Spinnmuster wird angewendet, um das begrenzte FoV des Sensors zu überwinden. Ein zweidimensionales Rauheitsmodell wird entwickelt, um die Merkmalspunkte für die Feinabstimmung und Registrierung der Punktwolke zu extrahieren. Zusätzlich setzt der Posenschätzer einen zeitlichen Bewegungsprädiktor ein, der dabei hilft, die Merkmalskorrespondenzen in der Karte für die schnelle Konvergenz des nichtlinearen Optimierers zu finden.

Wir haben die vorgeschlagenen Ego-Motion-Schätz- und Mapping-Methoden durch eine

Reihe umfangreicher Experimente validiert, die von Telepräsenz über handgetragene bis hin zu Drohnen-Setups reichen. Während der Experimente werden verschiedene Umgebungen erkundet, wie z. B. Innenlabore, Büros, Wohn- und Industriestandorte sowie eine breite Palette von gemischten Bedingungen. Die Ergebnisse zeigen, dass die Methoden in der Lage sind, hochpräzise Schätzungen mit geringer Latenz sowie Robustheit in Bezug auf schnelle Bewegung und Umweltzerstörung durchzuführen.

Acknowledgements

The work presented in this dissertation was carried out as a member of the academic staff at the Chair of Media Technology (LMT) at the Technical University of Munich. In spite of spending several years in the research domain of Robotics and AI, I am still thrilled and curious about its challenges. Getting to work in this field gives me the opportunity to participate in shaping the future and building a better world. This is mainly because I was fortunate enough to work with so many talented and visionary people whose enthusiasm motivated me while we were collaborating. I have been supported by several people, professionally and personally, during the past several years.

First of all, I would like to express my deep gratitude to my supervisor Prof. Eckehard Steinbach. For me, it has been an honor and a privilege to be a part of his team. He provided me with endless support and valuable feedback that helped shape my final dissertation. His professional attitude and enthusiasm were contagious and inspiring. In addition, I am grateful to Prof. Joachim Horn for agreeing to become the second examiner and to Prof. Norbert Hanik for chairing the thesis committee.

Many thanks go also to our project partners Lufthansa Technik AG, Helmut-Schmidt-Universität, and Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt for their support during our interesting and fruitful collaborations. I would like to thank the German Aerospace Center (DLR), and the Federal Ministry of Economic Affairs and Energy (BMWi) for supporting this research project.

Additionally, I would like to express my sincere appreciation to my former and current colleagues at LMT. In addition to friendships, the group has been a source of good advice as well as opportunities to collaborate. Their continued support has not just helped me professionally but has also made my time at LMT special and enjoyable on a personal level. For helping me begin my research with their experience and knowledge, I would particularly like to thank Dr. Tamay Aykut and Jingyi Xu. I would also like to thank Martin Oelsch for his fundamental contribution to the success of the Ki-Drohne project that we designed from scratch and developed during many sleepless nights. I would like to take this opportunity to extend my sincere thanks to Edwin Babaians for all his help and support throughout the course of this research. We had very fruitful exchanges, and I gained a lot of knowledge from him.

I am thankful for having such amicable colleagues and students at LMT, but I also feel the need to mention a few individuals in particular. Particular thanks go to Dr. Rahul Chaudhari, Dr. Dominik van Opdenbosch, Dr. Christoph Bachhuber, Michael Adam, and Martin

Contents

Abbreviations

Symbols

$\bar{\mu}$ mean of μ, defined on page. 14

2-DoF Two Degrees of Freedom. 69

2D Two Dimensional Space. 11, 28, 82

3-DoF Three Degrees of Freedom. 10, 11, 13

3D Three Dimensional Space. 2–4, 10, 11, 19, 20, 22, 23, 26, 27, 82, 102

6D Six Dimensional Space. iii, ix, x, 1–7, 25–30, 69, 81, 82, 101, 102

9-DoF Nine Degrees of Freedom. 11, 13, 82

A

AHRS Attitude and Heading Reference System. ix, 4, 10, 11, 27, 28, 31, 43

ANN Artificial Neural Network. 14, 15

ARS Angular Rate Sensors. 9

AVS Angular Velocity Sensors. 9, 10

C

CNN Convolutional Neural Networks. 14

CSPC Continuously Slicing Point Cloud. 30

D

DGPS Differential GPS. 3, 101

DL Deep Learning. 14

DNN Deep Neural Network. 14–16, 82, 101

E

EKF Extended Kalman Filter. ix, 12–14, 19, 29, 69

F

FOV Field of View. iii, 4, 23, 26–28, 82

G

GLONASS Global Navigation Satellite System. 7

GNSS Global Navigation Satellite System. 7, 8

GPS Global Positioning System. xi, 1–3, 7, 8, 25, 82, 101

I

ICP Iterative Closest Point. 20, 21, 102

IMU Inertial Measurement Unit. ix, 10, 11, 27, 28

INS Inertial Navigation System. ix, 1–3, 10, 101

L

LiDAR Light Detection and Ranging. iii, ix, x, 2, 4–7, 11, 19, 20, 23–30, 69, 70, 82, 83, 101, 102

LOAM LiDAR Odometry and Mapping. x, xiii, 22–24, 28, 102

LSTM Long-Short Term Memory. ix, 16–18, 81, 102

M

MARG Magnetic, Angular Rate, and Gravity. 10

MEMS Micro-electromechanical System. 8, 11, 12, 31, 43, 81, 82

MIMU Magnetic and Inertial Measurement Unit. ix, x, 2, 4, 7, 10–13, 15, 17, 20, 25, 27–30, 81, 82

ML Machine Learning. 14

N

NN Neural Network. 15, 101

R

RGB-D Red-Green-Blue-Depth. 8

RNN Recurrent Neural Networks. iii, ix, 5, 14–16, 18, 43, 81, 101

ROS Robot Operating System. 30

RPM Rounds Per Minute. 29

S

SLAM Simultaneous Localization and Mapping. iii, ix, x, 1, 2, 4–7, 18–23, 25–27, 29, 30, 43, 69, 81, 82, 101, 102

surfel Surfel Element. 21, 22, 102

U

UAV Unmanned Aerial Vehicle. 1, 6–8, 11, 25–27, 69, 70, 101, 102

UKF Unscented Kalman Filter. 12, 13

UWB Ultra Wideband. 8

V

V-LOAM Visual-LOAM. 23

W

WiFi Wireless Fidelity. 2, 8

Chapter 1

Introduction

1.1 Motivation

Figure 1.1: Airplane autonomous inspection using UAVs. An example where real-time motion estimates and mapping are needed simultaneously for navigation. In such applications, high-accuracy GPS/INS systems are impractical due to the weight, cost, and external magnetic perturbations. GPS signals are also unreliable underneath the structures of the hangar. Camera-based localization systems, on the other hand, provide poor performance and this is mainly due to the texture-less surfaces and far-field view. The UAV's navigation stack requires 6D motion estimation. The estimation latency is critical in such applications as the robot experiences a very fast motion behavior.

The problem of real-time Simultaneous Localization and Mapping (SLAM) is one of the most fundamental challenges in robotics. Localization and perception techniques for autonomous robot missions (see Figure. 1.1 for an example) have become one of the major drivers for the development of SLAM solutions in recent years. The mass production of

inexpensive, versatile, and compact Light Detection and Ranging (LiDAR) and Magnetic and Inertial Measurement Unit (MIMU) sensors, initially targeted at the smartphone and autonomous car markets, now enables autonomous robots to be equipped with such sophisticated sensor systems.

Robotic systems using multi-sensor SLAM can explore the environment and create a map while localizing themselves simultaneously. This problem is also sometimes referred to as the kidnapped robot problem [13], since the robot needs to determine its location and surroundings on its own without any additional infrastructure. In outdoor environments, the satellite-based Global Positioning System (GPS) provides information about absolute position, but not about the environment's structure, which is key to avoiding collision and performing path planning. There are, however, a number of possible application scenarios, such as urban canyons and indoor environments, which may require additional costly infrastructure for localization. Several such scenarios are classified as GPS-denied environments. A wide variety of approaches are available ranging from relatively inexpensive WiFi-based approaches to more expensive radar methods that provide centimeter-level accuracy. The downside is that many of these methods do not provide information about the environment and obstacles.

1.2   Towards Real-time 6D SLAM for Indoor Applications

It is a common requirement across a wide range of applications to create accurate Three Dimensional Space (3D) representations of the world and to localize the robot within these 3D representations (see Fig. 1.2a for an example). State-of-the-art technology in this field comes from the use of laser scanning. Lasers are especially appropriate for 3D reconstruction because they provide range accuracy that can be as low as a few centimeters regardless of the distance measured. Additionally, ambient light and surface normals are not significant factors. High spatial resolution can be obtained by using coherent laser beams. In addition, the high sample rate of laser measurements (up to millions of measurements per second) translates into excellent temporal resolution as well. The registration of LiDAR data is an issue in this regard. In order to obtain accurate reconstructions, the LiDAR position must be known with high precision during the measurements. Two methods are used in industrial reconstruction to achieve the required level of precision.

First, a mechanism stays stationary except for the movement of the actuator directing the laser beam (see Fig. 1.2b for an example). Data from several sources are collected and merged offline. A limited number of scans can be collected using this method, which is time-consuming and easy to obstruct the result.

An alternative approach is to use independent positioning systems (such as industry-grade GPS/INS systems) to register moving LiDAR data (see Fig. 1.2c for an example). Due to the weight, cost, or limited availability of GPS signals, high accuracy GPS/INS systems are often impractical. An example is shown in Fig. 1.2d. In this figure, an autonomous drone flies around an airplane for mapping and inspection. High-accuracy GPS/INS systems are too heavy to carry due to the limited flight payload. Additionally, in such applications, GPS

(a)   (b)   (c)   (d)

Figure 1.2: (a) A map of an indoor area. (b) A Leica scanner attached to a tripod. (c) A laser scanner and industry grade GPS/INS mounted on an autonomous car. The car is equipped with multiple GPS/INS systems and uses the Differential GPS (DGPS) to provide precise motion estimates. (d) Autonomous drone for airplane inspection and mapping. In such application, high-accuracy GPS/INS systems are impractical due to the weight and cost. GPS signals are also unreliable inside the hangar or below structures of the buildings. The latency of the estimation is critical for real-time navigation.

signals are unreliable underneath the structures of the hangar. Moreover, due to a large amount of external ferromagnetic material in such environment, INS data is also unreliable. This is mainly because the INS sensors are sensitive to external magnetic perturbations. On the other hand, cameras can capture visual patterns from their surroundings. The global shutter camera eliminates the difficulty of estimating the pose state as pixels in a picture share the same timestamp. However, cameras are sensitive to changes in lighting conditions. Furthermore, a homogeneously colored surface, for instance, can pose a challenge to methods that are based on vision.

The focus of this thesis is on the capability to localize a robot in real-time while creating high-precision maps from a moving platform without precise and independent positioning information being available. A common understanding is that to obtain such maps without independent positioning, localization and mapping must be solved simultaneously. However, the challenges include how to achieve this accurately in 6D, in real-time, robustly, with low latency, and in a small form factor.

This thesis seeks to combine range, inertial, and magnetic sensing to compute the ego-motion and reconstruct the environment in 3D. Even though it is possible to further enhance 6D reconstruction by using loop closure, this dissertation will focus on low-latency and real-time methods capable of providing high precision 6D motion estimation. The low-latency and real-time attributes make the resulting system suitable for autonomous navigation: the robust 6D motion estimation is critical for robot control, and the maps are crucial for obstacle avoidance and path planning. Further, the complete map of the environment can be used as an input for further processing such as scene segmentation, 3D reasoning, and virtual reality applications.

1.3 Principal Contributions

The thesis proposes to solve a low-latency state estimation problem by combining data from various sensors in order to provide 6D ego-motion in real-time and reconstruct the environment in 3D. In order to solve the aforementioned problem, several difficulties must be addressed. As part of the proposed research, different sensor modalities will be used, while sensors of each kind must also be characterized in distinct ways.

The range measurement capabilities of LiDARs are excellent. It is critical to keep in mind that LiDAR measurements are made continuously over time, resulting in a unique timestamp associated with each laser point. To create a coherent map from the laser points when the LiDAR moves with 6D motion, high-frequency motion estimates must be made to reduce the distortion caused by the LiDAR's intrinsic motion. The continuous motion over time is often approximated by a large number of variables in state estimation. Continuous slicing of incoming laser measurements in a multi-threaded pipeline is utilized to overcome this challenge. Additionally, the state-of-the-art laser scanners come with a limited Field of View (FOV). Although for 3D planar pose estimation these sensors are practical, this limitation results in poor 6D pose estimation, specially when encountering vertical and rotational movements. An actuated LiDAR system is utilized to overcome this limitation.

The ability of 6D pose estimation using the aforementioned LiDAR solutions crucially depends on the reliable estimation of attitude and heading angles in real-time [10], [11]. This additional source of sensory data can be attained mainly from an Attitude and Heading Reference System (AHRS). Robust attitude and heading estimation in an indoor environment with respect to a known reference is an essential component for 6D pose estimation. Affordable AHRSs are typically using low-cost solid-state MEMS-based sensors. The precision of heading estimation on such a system is typically degraded due to the encountered drift from the gyro measurements and distortions on the magnetic field sensing. In this regard, multi-sensor MIMU fusion is utilized to perform robust heading estimation with the ability to compensate for the external magnetic field anomalies.

This thesis explores a new method of fusing information from various sensors. As opposed to composing all sensor data into a single, full-blown problem, we divide the processing into a series of modules and solve them sequentially. Within each module in the system, an advantage of each sensor is explored, a sub-problem is solved, and results are generated for the following modules to process. The modularized processing pipeline also improves robustness, by selecting "healthy" sensor modes when forming the final solution. Here, "healthy" means the sensor data contains sufficient information to carry out the desired estimation. For example, when a LiDAR experiences a fast rotational movement, useful information becomes sparse causing estimation failures. Similarly, in an asymmetric or extruded environment such as a long and straight corridor, traditional solutions can produce ambiguous poses sliding along the corridor. Likewise, the AHRSs cannot estimate the attitude and heading angles precisely in the presence of external magnetic perturbations. In these cases, the proposed method determines a degraded subspace in the problem state space. During degradation, part of the problem is resolved within the subspace that is well-conditioned. Consequently, only healthy elements are combined in the final solution.

Scope of the thesis: In this thesis, a multi-sensor fusion technique that implements low-latency, real-time 6D SLAM for indoor applications is presented, which significantly enhances the performance of existing SLAM frameworks.
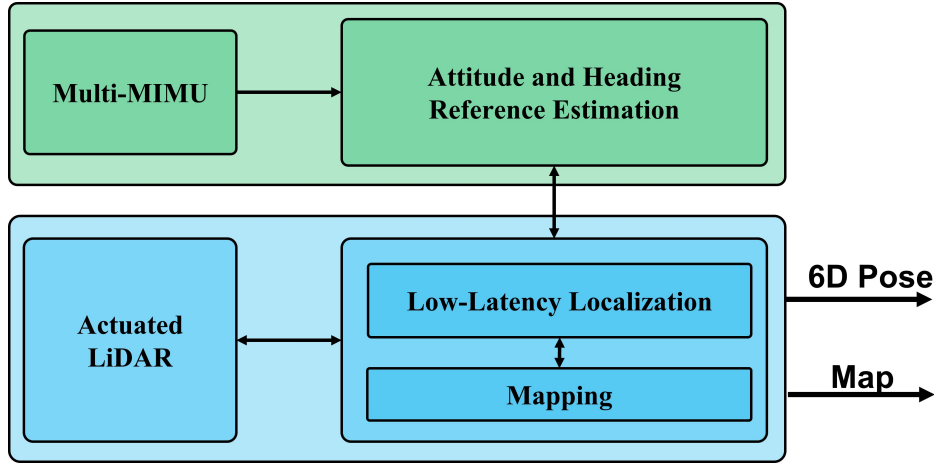
Figure 1.3: System overview of the proposed low-latency, real-time 6D SLAM system for indoor applications. We propose to tackle the problem by leveraging range, magnetic, and inertial sensing in a coarse-to-fine manner. This work is divided into two main subsections: robust attitude and heading estimation in an indoor environment using a multi-sensor fusion approach (green part), and low-latency 6D ego-motion estimation and mapping techniques with LiDAR-based systems (blue part).

Summary of Contributions: To solve the low-latency, real-time 6D ego-motion estimation and mapping problem, the thesis makes the following contributions:

- A skewed redundant magnetic and inertial sensor platform and the methodology to configure redundant sensing kinematic for non-orthogonal observation is developed. Additionally, a correlation-based preprocessing filter model is designed. Multiple RNN-based deep fusion models are investigated to improve accuracy and mitigate the external perturbation for robust heading estimation in a saturated indoor environment.

- The proposed processing pipeline is dynamically reconfigurable, which fully or partially bypasses failure modules and combines the rest system to deal with sensor degradation. It can handle environmental degradation and aggressive motion.

- A method embedded in the processing pipeline determines sensor degradation. It separates well-conditioned subspace from degenerate subspace in the underlying problem state space, solves the problem only in the well-conditioned subspace, and therefore actively eliminates sensor noise in the degraded subspace.

- A modularized processing pipeline to fuse data from range, magnetic, and inertial sensors for 6D ego-motion estimation and mapping. The resulting system achieves high accuracy and low drift. In LiDAR processing module, the method uses a voxel representation to efficiently retrieve registered laser points for scan matching.

- Also in the LiDAR processing module, the method conducts scan matching on multiple threads. Each thread processes a subgroup of scan slices while multiple subgroups are processed in parallel to carry out the onboard processing in real-time, which increases the update rate of the 6D pose estimation, and reduces the overall estimation latency.

- A novel sensor payload design in which a small, lightweight multi-line 360° laser scanner is actuated with a Lissajous pattern. This structure is employed to reduce the revisit time and provides a dense point cloud with consistency in scanning the environment. The designed platform meets the size and payload restrictions of mobile robot applications such as UAVs.

1.4   Thesis Structure

Chapter 2 introduces the theoretical background needed to understand the proposed concepts and the decisions made within the scope of this work. The proposed methodology is confronted with related work. The main advantages are discussed and compared to the state-of-the-art. **Chapter 3** presents a novel approach for robust indoor heading estimation based on skewed redundant magnetometer-only fusion. A correlation-based fusion technique is introduced for perturbation mitigation using the proposed skewed redundant configuration and the experimental results are presented. **Chapter 4** covers the deep-learning-based fusion of a skewed redundant magnetic and inertial sensor for robust attitude and heading state estimation in a saturated indoor environment and the experimental result of the proposed system is comprehensively discussed. **Chapter 5** introduces the techniques for low-latency SLAM based on LiDAR scan slicing and concurrent matching. The pipeline is developed for 6D ego-motion estimation with high update rate and low latency and the experimental results are discussed. **Chapter 6** concludes this work by summarizing the most distinct outcomes of this manuscript, discussing their limitations, and briefly outlining some potential conceptual improvements for future work.

Parts of this thesis have been published in international peer-reviewed scientific journals [1]–[4], and conferences [6]–[9].

Chapter 2

Fundamentals and Related Work

This chapter introduces the relevant background needed to better understand the proposed methodologies within the scope of this work. Prior art that technically relates to this work is surveyed and discussed. Section 2.1 presents the basics of sensor systems and covers topics from external reference systems to LiDAR acquisition systems. Assets and drawbacks thereof are briefly debated to better comprehend the strategical decisions made for this work. Section 2.2 deals with state-of-the-art methodologies that aim to solve attitude and heading reference estimation using multi sensor fusion approaches. Section 2.3 deals with state-of-the-art methodologies for SLAM. The last section delves into the area of low-latency 6D SLAM using LiDARs and MIMUs and highlights the basic fundamentals used in this work.

2.1 Sensor Systems

In recent years, multi-rotor UAVs have gained popularity as improvements in battery efficiency and computational power, as well as sensor size, made payload limitations less relevant. Additionally, they offer great mobility with six degrees of freedom and can hover, take off, and land vertically. It is, however, an unstable system that requires constant effort to maintain a controllable state in the air and therefore requires sensors to determine its position at a high frequency to correct for external influences, such as airflow. Further, for the inspection task, the UAV delivers the inspection equipment to the measurement spots while avoiding collisions with the surrounding area. It is therefore of primary importance to determine the pose of the drone as it could otherwise cause damage to itself and the environment. In order to increase the system's accuracy and prevent failures, multiple sensors are therefore used. Multi-sensor systems have been proven capable of autonomous flight [12] [13], as well as inspection operations [14]. In this section, we review some of the common sensors used in SLAM applications and refer the reader to Brena et al. for a comprehensive survey of potential sensors [15].

2.1.1 External Reference Systems

A Global Navigation Satellite System (GNSS) like Global Positioning System (GPS) or GLONASS uses the triangulation procedure in order to determine a point on the earth by gathering information from at least four satellites. The accuracy ranges anywhere between a

few meters to several hundred meters. There are many advantages of using a GPS module, including its small size and low cost, which makes it useful to position commercial robots outdoors, and this is also true when there are external perturbations in sensory data. On the other hand, it should be noted that GNSSs are typically not available in the following circumstances: inside buildings, underwater, in suburban environments, or in other areas where satellites are not directly in sight. Consequently, it cannot be applied to the localization of mobile robots that are used in an indoor environment. In the meantime, there are also indoor localization systems with external references, such as Bluetooth, Wireless Fidelity (WiFi), or Ultra Wideband (UWB) [16]. A large number of researchers explored UWB [17] [18] [19] as well as Bluetooth [20] [21] for indoor applications. The problem with these systems is that they are less flexible because they require beacons with a known position that must be installed and are not suitable for adaptable indoor settings.

2.1.2 Vision Sensors

During the recording of images from a camera, visible characteristics of the environment can be captured. With the help of these images, one is able to produce maps of the surroundings and, with the assistance of these maps, one is able to locate the target. In spite of the fact that a single picture does not contain any information about position, it is derived by comparing multiple pictures and by measuring the difference in the location of landmarks, like trees or houses. Using stereo cameras, as there is no temporal difference between the images, is an excellent method for estimating distance. Red-Green-Blue-Depth (RGB-D) cameras, on the other hand, can supplement the images with a measure of distance. The only limiting factor is the computational overhead of low-cost cameras. Additionally, in order to properly perform a visual navigation task, adequate ambient light and a sufficient number of visual features in each and every image are essential. It is worth mentioning that vision-based navigation systems were the first to produce dense maps for UAV applications and they remain an active area of research and development today [22]. Several recent achievements have come from monocular [23] [24], stereo [25] [26] and RGB-D [27] [28] approaches, and Lu et al. provide an overview [29].

2.1.3 Environmental, Magnetic and Inertial Measurement Unit

2.1.3.1 Environmental Sensors

Environmental sensors include barometric pressure sensors that measure air pressure [30], relative humidity sensors measuring relative humidity levels, and temperature sensors that measure ambient temperature to enable adaptive multi-sensor fusion solutions [31] [32].

2.1.3.2 Magnetometer Sensors

Magnetometer sensors detect the magnitude of magnetism and geomagnetism generated by a magnet or current. In this context, Micro-electromechanical System (MEMS) variants are widely used into aerial applications for their light weight and low cost, making them a suitable candidate for being utilized by UAVs [33]. Considering that the magnetic observation is

a vector field, it is characterized by its magnitude and direction or its components along the coordinate axes. A sensor with at least three axes is required to determine the Earth's magnetic field vector. In a Cartesian coordinate system at the measurement point p, the magnetic field vector $\mathbf{b}_p \in \mathbb{R}^3$ can be defined as

$$\mathbf{b}_p = b_{\mathrm{p},x}\,\boldsymbol{i} + b_{p,y}\,\boldsymbol{j} + b_{p,z}\,\boldsymbol{k}\,, \tag{2.1}$$

where the x-axis is oriented along the geographical meridian, and the direction to the north is positive, the y-axis along the parallel with a positive direction toward the east, and the z-axis is directed downward. The observation point p is the origin of the measurement coordinate system.

The magnetic field of a dipole, similarly, is characterized by its magnetic moment $\mathbf{M}$. The strength and direction of this field depend not only on the magnitude of the magnetic moment but also on its orientation relative to the direction of the magnetic field [34]. In the presence of multiple dipoles, the magnetic field $\mathbf{M}_g$ is defined by the sum of all magnetic moments $\mathbf{M}_i$ and is given by

$$\mathbf{M}_g = \sum \mathbf{M}_i\,. \tag{2.2}$$

The curl of the magnetic vector potential, the magnetic field $\mathbf{b}_p$, is governed by the density of magnetic moments at the measurement point and is derived classically using the Biot-Savart law [35]. The combined field can change the measurement field vector $\mathbf{b}_p$ due to the presence of multiple dipoles. This change in the magnetic field is the leading cause of heading errors in an indoor environment [36]. To determine the Earth's true magnetic field, we consider the measured magnetic field as a combination of the external magnetic perturbation and the Earth's magnetic field. Taking Eq. 2.2 into consideration, the measured magnetic field $\mathbf{b}^Q \in \mathbb{R}^3$ in the sensor frame with center point p can be defined as

$$\mathbf{b}_p^Q = \mathbf{b}_p^E + \mathbf{b}_p^\epsilon\,, \tag{2.3}$$

where $\mathbf{b}_p^\epsilon$ is the external magnetic field disturbance and $\mathbf{b}_p^E$ is the Earth local magnetic field.

2.1.3.3 Inertial Sensors

Unlike other data sources, inertial sensors, such as accelerometers and gyroscopes, work independently of external input since they are based solely on natural phenomena. A single-axis accelerometer can measures proper acceleration [37]. The rate of change of velocity (the acceleration) of a body in its own instantaneous rest frame is known as the proper acceleration. It is important to understand that this is different from coordinate acceleration, which is acceleration in a fixed coordinate system. In the case of two or more accelerometers that are coordinated, they are capable of measuring differences in proper acceleration, especially gravity, across their separation - that is, the gradient of the gravitational field. Single- and multi-axis accelerometers can detect both the magnitude and the direction of the proper acceleration, as a vector quantity. Thus, multi-axis accelerometers allow the measurement of attitude orientation with respect to the gravitational field. In a similar vein, a gyroscope sensor is a device that can measure and maintain angular velocity and orientation. Gyroscope sensors are also referred to as Angular Rate Sensors (ARS) or Angular Velocity Sensors

(AVS). The ability to measure angular velocity in a 3D coordinate frame is achieved by using two or more gyroscopes that operate in a coordinated manner.

Similar to the magnetic field vectors, accelerometer and gyroscope data can be measured in a defined sensor coordinate frame p. Accordingly, the observed acceleration $\mathbf{a}_p \in \mathbb{R}^3$ in the body frame with center point p can be defined by

$$\mathbf{a}_p = a_{p,x}\,\boldsymbol{i} + a_{p,y}\,\boldsymbol{j} + a_{p,z}\,\boldsymbol{k}\,, \tag{2.4}$$

where $\mathbf{a}_p$ is the acceleration vector in m/s^2. Likewise, the angular rotation $\mathbf{v}_p \in \mathbb{R}^3$ in the body frame with center point p is defined by

$$\mathbf{v}_p = v_{p,x}\,\boldsymbol{i} + v_{p,y}\,\boldsymbol{j} + v_{p,z}\,\boldsymbol{k}\,, \tag{2.5}$$

where $\mathbf{v}_p$ is the angular velocity vector in $rad \cdot s^{-1}$.

2.1.3.4   Inertial Measurement Unit (IMU)

An IMU is a system that measures and reports a body's specific force, angular rate, and the orientation of the body, using a combination of 3-DoF accelerometers and 3-DoF gyroscopes [38].

2.1.3.5   Magnetic and Inertial Measurement Unit (MIMU)

An MIMU uses Three Degrees of Freedom (3-DoF) accelerometers, 3-DoF gyroscopes, and 3-DoF magnetometers, to continuously calculate the attitude and heading orientations. Typically, the orientation estimation using MIMU considers no external references [38] [39]. This means the orientation estimations are based on the starting point.

2.1.3.6   Attitude and Heading Reference System (AHRS)

An AHRS consists of sensors on three axes that provide attitude and heading information, including roll, pitch, and yaw. It is also referred to as Magnetic, Angular Rate, and Gravity (MARG) sensors and consist of 3-DoF gyroscopes, 3-DoF accelerometers and 3-DoF magnetometers. The main difference between the MIMU and the AHRS is that the AHRS provides the orientation estimation with respect to a know reference [32] [36]. The known reference in this context means the estimated orientations are with respect to the Earth's geometrical characteristics. The known reference of AHRS is mainly comes from the Earth's gravitational and magnetic field for attitude and heading angles, respectively.

2.1.3.7   Inertial Navigation System (INS)

An INS uses 3-DoF accelerometers, 3-DoF gyroscopes, 3-DoF magnetometers, and barometric pressure sensors to continuously calculate position, orientation, and velocity of a moving object without the need for external references [40] [41].

2.1.4 **Light Detection and Ranging (LiDAR)**

Light Detection and Ranging (LiDAR) is a system that uses a laser to determine distances by aiming the laser at an object or a surface. LiDAR determines distance by measuring the time it takes for the reflected light to return to a receiver. Additionally, the light intensity which is measured from reflected laser beam is primarily used to produce digital ono-colored images of the surrounding environment. Hence, they are independent of outside light sources. Nevertheless, because of the active light-emitting process, they also have a higher power consumption as compared to the other presented sensors, and in addition are more expensive. As in modern LiDAR systems, the laser scanner is rotated around an axis to obtain a 360-degree panoramic perspective of the environment. It is possible for some LiDAR systems to perform over 1 million measurements per second by combining multiple laser scanners pointing in different directions [42]. There is therefore the possibility of building detailed maps only from LiDAR data, even though cameras are able to record a lot more data in the same time period as LiDAR. There are a number of Two Dimensional Space (2D) LiDARs that are typically found in UAVs due to the payload constraints [43] [44] [45], but 3D LiDARs are also quite common [46].

2.2 **Attitude and Heading Reference Estimation Using Multi-MIMU**

Attitude and heading estimation is one of the fundamental necessities for human-computer interaction, robot control, as well as robot navigation [47], [48], [5]. The use of miniature inertial sensors has become quite common in robotic applications. A general Magnetic and Inertial Measurement Unit (MIMU) in this context, contains a 3-DoF accelerometer, 3-DoF gyroscope, and 3-DoF magnetometer. In contrast to IMUs or MIMUs, which generally provide only the raw sensory data, an advanced Attitude and Heading Reference System (AHRS) utilizes the Nine Degrees of Freedom (9-DoF) sensory data and applies the sensor fusion algorithm in order to provide absolute orientations. The roll, pitch, and yaw estimation are generally defined with respect to the earth's geometrical characteristics. The attitude and heading estimation specifies the angular measurements on the vertical and horizontal plane with respect to the local level frame and the geographical true north [49].

The main challenges in such sensors can be divided into multiple modalities; first is the accuracy of such estimation in a highly congested environment and in the presence of the external perturbations. The other modalities are the latency, update rate, and the reliability of the estimation. In order to increase the accuracy and the reliability of such sensors, the cost plays an important role. In this regard, multi-redundant low-cost platforms are well known in this context [6], [50], [51].

2.2.1 **Previous Work:**

Determining error-free and reliable heading angle with respect to a known reference is problematic in the case of indoor applications. This problem is mainly because of different sources of errors in the MEMS-based magnetometer measurements [52]–[55]. The most dominant

source of error for such a magnetometer sensor is the presence of magnetic sources in an indoor environment. Permanent magnets and ferromagnetic materials are the preeminent cause of such magnetic perturbation. In this regard, calibration of the MEMS-based magnetometers for heading estimation has been investigated thoroughly during the last decades [52], [56]–[58]. Probabilistic models for sensor calibration and fusion such as the Extended Kalman Filter (EKF), the Unscented Kalman Filter (UKF), and Ellipsoid Fitting are well known in this field [59]–[62]. However, dynamic magnetic deviation caused by external anomalies cannot be fully compensated using the sensor calibration methods [63], [64]. To overcome the aforementioned problem, redundant sensor systems have shown promising results as they are using redundant observations to determine the source of error and increase the estimation accuracy [65]–[67].

Multiple magnetometers combined with gyroscopes were used to approximate the external magnetic disturbances and provide perturbation-free heading estimation [66], [68], [69]. In these works, the authors used an estimation of the expected magnetic field using the temporary rotation vector to compensate for the external magnetic perturbation. These approaches need a precise and simultaneous calibration of the magnetometers and the gyroscopes to be able to provide an acceptable heading estimation. In [70], a magnetic heading compensation method for disturbances introduced by near-surface interference was proposed. The authors used inverse magnetic anomaly vector estimation to compensate for the external perturbation. The main limitations of this approach are the predefined magnetic field model and one-dimensional perturbation mitigation. In general, these approaches are usually bulky, expensive, and often rely on expensive calibration procedures.

Nonetheless, it has been shown that the accuracy of the magnetic field measurement is significantly improved by using redundant magnetometer sensors [71]–[74]. However, the non-deterministic errors cannot be completely removed in real application scenarios due to their transient nature. This is mainly because of limitations in the system identification procedure, non-linear dynamic errors, and the use of strict calibration models. The skewed redundant inertial sensor fusion, on the other hand, provides a more reliable output, as shown in [51], [75]–[77]. The skewed configuration for redundant sensor geometry has been widely investigated for attitude estimation by applying this method on the observations from inertial-only sensors (e.g., accelerometers and gyroscopes) [41], [78]. However, these methods have been poorly investigated for redundant magnetometers.

2.2.2   MIMU Errors

MIMU or Multi-MIMU systems are mainly used to track an object, so that the tracking system is able to determine the object's current position and/or orientation by analyzing the sensor data. Sensor information may be corrupted by a variety of errors, and then, during the fusion, the errors will accumulate, leading to the phenomenon known as drift error. In order to prevent this, errors in the collected data should be minimized or compensated for. MIMU errors can be classified as deterministic or stochastic based on their nature [79], [80]. For instance, misalignment errors and scale factor errors fall under deterministic error types. In contrast, stochastic errors are primarily characterized by noise [81]. The bias in sensor

data usually consists of multiple components, some of which are stochastic, some of which are deterministic [82]. As a rule of thumb, all of the errors have an impact on the output of a sensor (see Figure 2.1).
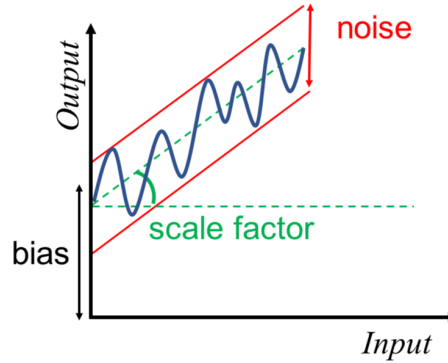

Figure 2.1: Magnetic and inertial sensors contain two main types of errors which are deterministic errors like scale factor, bias, misalignment and stochastic errors such as bias instability and scale factor instability. Deterministic errors are the main part of error compensation algorithms. The bias in magnetometer sensor data usually consists of multiple components, some of which are stochastic, some of which are deterministic [82].

2.2.3   Skewed Redundant MIMU Fusion

A combination of a 3-DoF gyroscope, a 3-DoF accelerometer, and a 3-DoF magnetometer is typically used in low-cost sensor platforms. The estimated angles are defined as the corresponding Euler angles roll (α), pitch (β), and yaw (γ). In order to improve the robustness and the accuracy of the attitude and heading estimation, the measurements provided by the gyroscope, the accelerometer, and the magnetometer are typically fused (e.g., Extended Kalman Filter (EKF) or Unscented Kalman Filter (UKF)). As previously discussed, multi-MIMU provides more robust estimation in comparison to a single-MIMU fusion. Typically, in such a multi-MIMU fusion, sensors are mounted in an array configuration [83], [84]. However, a significant outcome from [50], [51] shows that an ideal setup of redundant inertial sensors in terms of fault tolerance and noise resistance is a "skewed" configuration. In the skewed-redundant configuration, the delicate axes of each triaxial sensor are not aligned, and transformations are performed to align them in a virtual coordinate frame. This method has been examined on different sensors such as accelerometers, and gyroscopes and showed an improvement in the accuracy and the perturbation mitigation [51], [75], [77]. Motivated by the aforementioned works, within the scope of this thesis, we propose a skewed-redundant tetrahedron platform in order to increase the reliability and the accuracy of the attitude and heading estimation. In this work, we employed triple single chip 9-DoF MIMU sensor from Bosch-Hillcrest Labs.

2.2.4 **Extended Kalman Filter (EKF)**

In practice, the assumption that the state transition and measurement are linear with a Gaussian distribution is rarely true. Using the Extended Kalman Filter (EKF) one can overcome the nonlinearity problem by linearizing the state transition function as well as the measurement function. In the EKF the state transition and measurement are represented as functions g and h with additive Gaussian noises η_t and δ_t:

$$x_t = g(u_t, x_{t-1}) + \eta_t \tag{2.6}$$

$$z_t = h(x_t) + \delta_t \tag{2.7}$$

with:

$$g'(u_t, x_{t-1}) := \frac{\partial g(u_t, x_{t-1})}{\partial x_{t-1}} \tag{2.8}$$

$$h'(x_t) := \frac{\partial h(x_t)}{\partial x_t} \tag{2.9}$$

one can define the Jacobian matrices J_t around the mean μ_{t-1} ($\overline{\mu}$), and H_t around $\overline{\mu}_t$ as:

$$g'(u_t, \mu_{t-1}) := J_t \quad \text{and} \quad h'(\overline{\mu}_t) := H_t \tag{2.10}$$

The EKF algorithm is in many ways similar to Kalman Filter with the difference of linearization of the observation and motion functions. More details to EKF can be also found in [85].

2.2.5 **Deep Sensor Fusion**

The term Deep Learning (DL) refers to a subclass of Machine Learning (ML) that is able to solve more complex problems generally due to a comprehensive architecture of neurons that enables them to learn more sophisticated relations between inputs and outputs. Developing deep learning models requires the use of powerful processing units that can handle many layers of underlying data. In contrast to machine learning algorithms that require human intervention in order to perform feature engineering, Deep Learning algorithms are capable of mapping complex functions from the input to the output without requiring user intervention. Deep Learning can also be described as Hierarchical Learning or Structured Learning. The mechanism of Deep Learning is a Artificial Neural Network (ANN) with multiple hidden layers, often referred to as a Deep Neural Network (DNN). A regular ANN and a DNN are shown schematically in Figure 2.2.

Using a hierarchy of layers in DNNs allows for the development of complex relationships between inputs and outputs. The aim of the network is to identify high-level features that are perceived by low-level properties. As an example, deep learning has been successfully applied to image and text classification, natural language processing, autonomous driving, weather and exchange rate forecasts, human gesture prediction, etc. Currently, there are two types of DNN that are most commonly used: Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN). The CNN models are used quite often when it comes to computer vision applications. In the meantime, RNNs are commonly employed for the analysis of time-series data as well as for signal processing applications. Due to this work
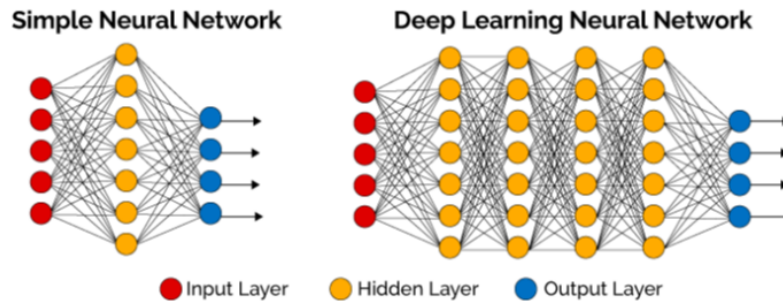
Figure 2.2: The difference between a standard Neural Network (NN) and a Deep Neural Network (DNN). While Deep Learning incorporates neural networks within its architecture, there's a stark difference between DNN and NN. DNN leverages a series of nonlinear processing units comprising multiple layers for feature transformation and extraction. It has several layers of artificial neural networks that carry out the machine learning process [86].

focusing on sensor fusion applications, the RNN models are addressed more deeply in this section to provide further insights.

2.2.5.1 Recurrent Neural Networks (RNN)

Standard artificial neural networks have no memory of previous states, except the last one: they start at the state in which they were trained. As a result, after the initial training is completed, the network does not exhibit much evolution. ANNs can only learn from data currently presented to them, which makes it useful only for independent inputs. The problem is that this is not true for all input sequences, so a type of neural network was developed to help overcome it - Recurrent Neural Networks (RNN)s. Figure 2.3 (left) illustrates what distinguishes RNNs from other types of neural networks by showing how they have loops inside themselves that allow neurons to remember information [87]. In contrast to ANNs that perform based on the current inputs, RNNs are based on series of inputs and previous states. Figure 2.3 (left) depicts an RNN consisting of a basic unit of one neuron with input Xt and output ht. Data are passed through several steps in the network by means of a loop. An unfolded version is illustrated in Figure 2.3 (right): each chain carries information to its successor.
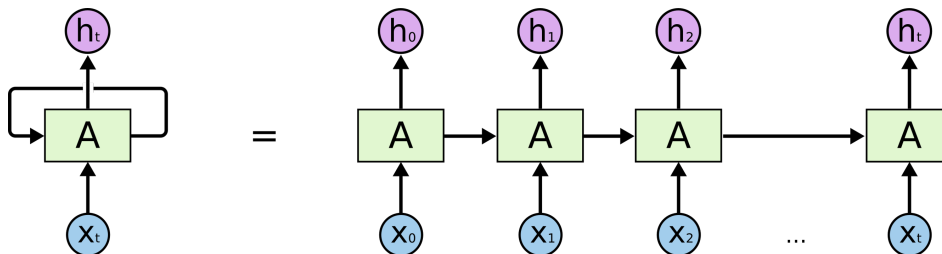

Figure 2.3: Compressed (left) and unfolded (right) basic recurrent neural network [86], [88]

As is depicted in Figure 2.3, the state of the unit at time t is also an input for the next step at time $t + 1$. In theory, this should mean that all the previous information is available at the new step; in practice, however, if there are many steps involved, the network exhibits

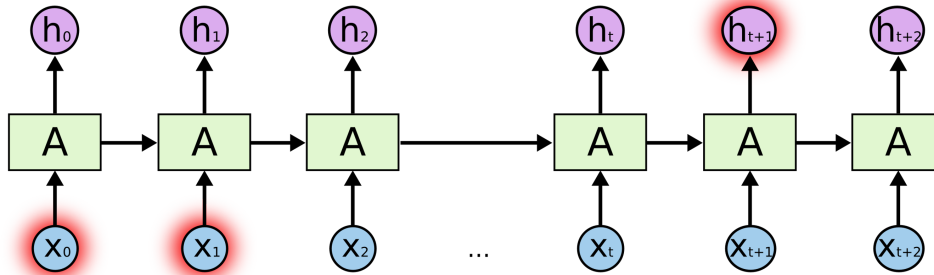vanishing gradient behavior [89], [90]. This long-term dependency problem is illustrated in
Figure 2.4.

Figure 2.4: Long-term dependency problem in Recurrent Neural Networks (RNN) networks. In theory, all the previous information must be available at the new step; in practice, however, if there are many steps involved, the network can not remember. This problem is called vanishing gradient problem [89], [90].

Due to the long-term dependency on the basic version of RNN, different modifications
have been proposed to overcome this issue. Long-Short Term Memory (LSTM), Bidirectional
LSTM, and Gated RNN are in this category. In this work, we have utilized LSTM-based
DNN models for sensor fusion application. In the following section, we dive even deeper
into LSTM-based models.

2.2.5.2   Long-Short Term Memory (LSTM)

Long-Short Term Memory, also known as LSTM, is one of the most popular choices of RNNs
when long-term dependencies are considered. This means that it can be used to solve the
vanishing gradient problem which can occur in sequence learning. LSTMs were originally
proposed by Hochreiter and Schmidhuber in 1997 and are still popular today [91]. LSTMs
are distinguished by the ability to retain certain information for a long period of time. Typically, a chain-like formation of modules repeats itself in an RNN, but in its most basic form,
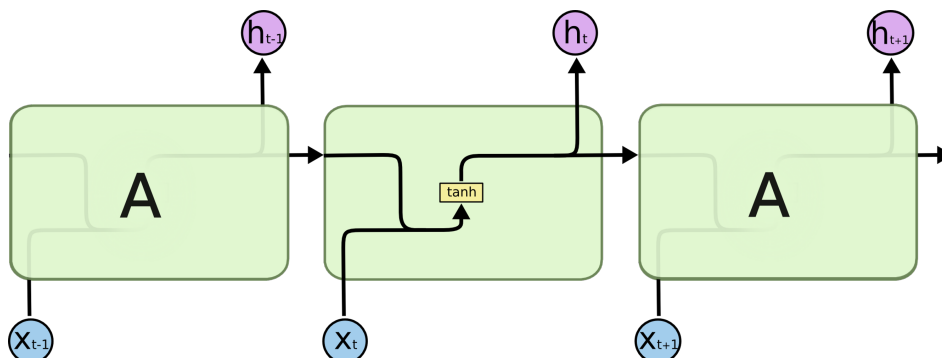the RNN can be composed of just one $tanh$ layer (see Figure 2.5).

Figure 2.5: Standard Recurrent Neural Networks (RNN) containing a single layer.

Together with $tanh$, sigmoid activation functions are used in LSTM. As explained in the
relevant illustration earlier, $tanh$ keeps the values in the range [1; 1], while sigmoid has a

range of [0; 1]. Using this strategy, one can learn which data should be kept and which should be discarded by multiplying the values by 0 if they need to be forgotten and by 1 otherwise.
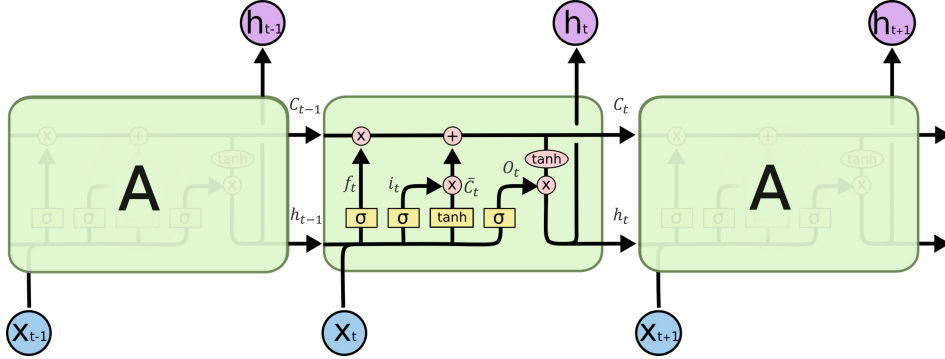

Figure 2.6: Overview of state-of-the-art LSTM cell architecture and a network of LSTM cells. A common LSTM unit is composed of a cell, an input gate, an output gate, and a forget gate. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell [91].

To implement this, LSTMs have gates of three types: input, forget, and output gates [92]. The cell architecture is presented in Figure 2.6. The gate equations for LSTMs are as follows. For the input gate, we can define i_t as

$$i_t = \sigma \left(w_i \left[h_{t-1}, x_t \right] + b_i \right),$$ (2.11)

and for the forget gate f_t we have

$$f_t = \sigma \left(w_f \left[h_{t-1}, x_t \right] + b_f \right),$$ (2.12)

while for the output gate the o_t is

$$o_t = \sigma \left(w_o \left[h_{t-1}, x_t \right] + b_o \right),$$ (2.13)

where $\sigma(\cdot)$ is the sigmoid function, $w.$ is the weight vector, h_{t1} is the output of the previous LSTM block at time step $t1$, x_t is the input at the current time step, and b is the bias. The equations describing the cell state, the candidate cell state, and the final output of the current cell are as follows. The cell state $\bar{c}_t$ is defined by

$$\bar{c}_t = \tanh \left(w_c \left[h_{t-1}, x_t \right] + b_c \right).$$ (2.14)

and the candidate cell state is

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \bar{c}_t$$ (2.15)

and the final output defined by

$$h_t = o_t \cdot \tanh(c_t).$$ (2.16)

This is a brief description of the standard LSTM node. The literature addresses a range of modification approaches, such as an LSTM with a peephole interface [93], where all the gate layers can read in to the current state. The other notable changes include depth gated RNNs [94] as well as clockwork RNNs [95].

2.3 Simultaneous Localization and Mapping (SLAM)

The main requirement of autonomous robot navigation systems is that the robot should be able to plan and execute its paths in an unknown environment without human assistance [96]. The process of route planning requires the robot to make decisions based on its surroundings, which has an impact on how it acts. This requires a real-time perception of the surrounding environment [97]. Having this awareness also permits the control system to receive position feedback, which is essential for the successful execution of the planned movement. This environment perception is achieved by using different types of sensors helping the moving platform position itself in reference to surrounding landmarks [98]. Therefore, the reason for performing localization and mapping is twofold; On one hand, the map itself may be the primary goal of the operation, i.e. it is intended to be used for reporting to a human correspondent, such as in the case of a construction site. On the other hand, with a detailed map, it is more effective to accurately localize the robot and plan the future path trajectory. The general problem in such is that the location and state of the robot are poorly known and can only be determined with noisy sensors [99]. Furthermore, localization and mapping are both associated with each other because localization methods require a map in order to refer to it as a reference for current measurements, while for building a map, the location, where the data was taken, is required in order to insert the measurements at the right spot on the map.

This problem is referred to as Simultaneous Localization and Mapping (SLAM). By using the SLAM technique, a robot can locate itself in reference to an initial start point while constructing a map of its surroundings [99], [100]. At each points of the map, landmarks have been inserted with the position and orientation of the robot at the timestamp when they were captured. The robot will use the most recent version of the map to plan its trajectory, taking into account the already registered objects [101], [102]. Two types of SLAM can be distinguished based on the sensors used to sense the environment: visual-inertial SLAM [103], [104] and Laser SLAM [105], [106].

Related Work: Initially, researchers discussed the topic of localization and mapping separately. With the addition of the works of Chatila and Laumond [107] and Smith and Cheeseman [108] SLAM evolved in the 1986 IEEE Robotics and Automation Conference in San Francisco, California [109]. In continuation, Cadena et al. divide the progress in the field into three eras: the classical era (1986 - 2004), the algorithmic-analysis era (2004 - 2015), and the robust-perception era (since 2016) [110]. A brief overview of the classical period is provided by Durrant-Whyte and Bailey [109] [111], when fundamental problems were defined and basic challenges were solved. Additionally, the probabilistic approach to SLAM with

EKF, Rao-Blackwellized particle filters, and maximum likelihood estimation was brought to light; Stachniss et al. explain these in great detail in [112]. After that, algorithmic analysis covered issues such as observability, consistency, and convergence for the SLAM problem [113]. At the same time, factor graphs were extensively studied, and open-source libraries for SLAM were developed [114]. From that point forward, the robust-perception age cover robustness and a high-level understanding of the environment, while being resource-aware and employing a task-driven perception [110]. In recent surveys, attention has mainly been focused on individual aspects of SLAM, like multi-robot [115], autonomous driving [116], or visual SLAM [117] [118]. Huang et al. provide a comprehensive overview of the latest algorithms [119].

2.3.1 Current Challenges

Despite Durrant-Whyte and Bailey's belief that SLAM had been solved on a conceptual level in 2006, it is still considered an active topic of research today [109]. The availability of sufficient solutions depends on a combination of factors, including the robot, the environment, and the performance requirements, such as accuracy or speed. We present our findings with the assistance of Cadena et al. [110] as the main challenges today are robustness, scalability, and latency. In robotic systems, failures are omnipresent, and having the ability to detect and recover from them is essential. Consequently, future SLAM systems will have to be fail-safe, failure-aware, and capable of compensating for hardware failures. While environments are often assumed to be static, they are very dynamic, not only for short periods of time like seconds or minutes but continuously. This calls for maps to be deformable and able to represent properties of the environment with adaptable descriptors, even when there are different levels of detail. Maps should be able to reflect the environment's properties in a semantically accurate manner too. Additionally, navigation operations require an increase in update rates, both for pose estimation and for maps. A map must therefore be able to store information over a long time span, and cover a large area, as well as be distributed among multiple robots [120], [121]. Latency, on the other hand, has become more and more crucial in recent years as the ability to support real-time navigation for fast systems, such as drones and autonomous cars, has become a necessity [122]–[124]. Finally, SLAM systems should also be adaptable to a wide range of robot applications, which means that the workspace and the dynamic of the system, as well as the environment, should also be considered.

2.3.2 LiDAR SLAM

In this section, we review methods for data association of the LiDAR sensor measurements as it is the primary input for localization in this work. We use a LiDAR system for SLAM instead of cameras. Despite the popularity of cameras, the visual feature distribution in an industrial environment such as a hangar is not sufficient for current state-of-the-art algorithms, like ORB-SLAM, to achieve the accuracy we require [26], [104]. A laser scanner performs successive range measurements, and by rotating it, it records a surrounding view. Based on the different scanning angles one can integrate multiple measurements into a 3D

point cloud. The accuracy of the point cloud, however, depends on the precise angle of the laser scanner during the measurement. While these angles are known with high precision, mobile robots move through the environment with an unknown motion. This motion causes the point cloud to be distorted, which is an inherent challenge to LiDAR SLAM. It has been shown that scanning skewing is a fundamental factor that can lead to incorrect rotation estimations. A common cause of this phenomenon is the motion of the platform when scanning and the fact that a single scan is made up of successively scanned points over a certain period of time [125]. Variable degrees of distortion can result from different types of motion. Non-rigid distortion is especially noticeable with rotational movement, affecting particularly points that are far from the sensor. Particularly, when the motion involves rotational and vertical translation simultaneously, the distortion must be compensated using an external MIMU sensor [126], [127].

In order to derive a relation between two point clouds taken in the same environment, the alignment of the clouds must be performed. This relation describes how two consecutively observed point clouds, or a position on a global point cloud, has changed with respect to each other, or with respect to the origin. Considering the point sets $\mathcal{P}_1$ and $\mathcal{P}_2$ in correspondence $\mathcal{P}_1 \leftrightarrow \mathcal{P}_2$, and that they are related via a rigid body transform, we seek to estimate $\boldsymbol{R}$ and $\boldsymbol{t}$. Let us define the homogeneous 3D affine transformation matrix $\boldsymbol{T} = (\boldsymbol{R}, \boldsymbol{t}) \in SE(3)$ as a rigid body transform composed of a rotation matrix $\boldsymbol{R} \in SO(3)$ and a translation vector $\boldsymbol{t} \in \mathbb{R}^3$, as

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{R} & \boldsymbol{t} \\ 0 & 1 \end{pmatrix}, \tag{2.17}$$

and accordingly, the transformation between the two point sets can be defined as

$$\mathcal{P}_{1,a} = \boldsymbol{R}\mathcal{P}_{2,b} + \boldsymbol{t} = {}^{a}\boldsymbol{T}_b \mathcal{P}_{2,b} \,, \tag{2.18}$$

where ${}^{a}\boldsymbol{T}_b$ is a transformation from $\mathcal{P}_{1,a}$ with the coordinate frame a to $\mathcal{P}_{2,b}$ with the coordinate frame b. The problem is well-studied with various closed-form solutions in the literature and in the following, we discuss methods for the alignment of point clouds and their association with maps.

2.3.3   Point-Cloud-based SLAM

Several algorithms align point clouds with an iterative approach, including Iterative Closest Point (ICP) [128], [129], which is the basis of many methods used today. We refer to Rusinkiewicz and Levoy for an early and Pomerleau et al. for a more recent review of variants [130] [131]. ICP alternates between a correspondence search and an optimization step. Correspondences are identified based on the nearest neighbor search and subsequently, the transformation is calculated by the minimization of an error metric describing the distance difference within the correspondences. The iterations are necessary since nearest neighbor searches are more likely to be able to find correct correspondences when the point clouds are closer aligned (see Figure 2.7). It is also important to keep in mind that due to the repeated searches for the nearest neighbor, the algorithm will only converge to a local minimum, so

a sufficient prior is necessary. This issue also can be address using coarse alignment techniques. In this regard, Al-Nuaimi outlines several approaches for coarse alignment methods in his work [132]. However, to form a prior for SLAM application, it is assumed that the previously determined position from the laser is sufficient enough. While the original algorithm directly operates on points, a point-to-plane approach, with planes based on the point normals, converges quicker [130].
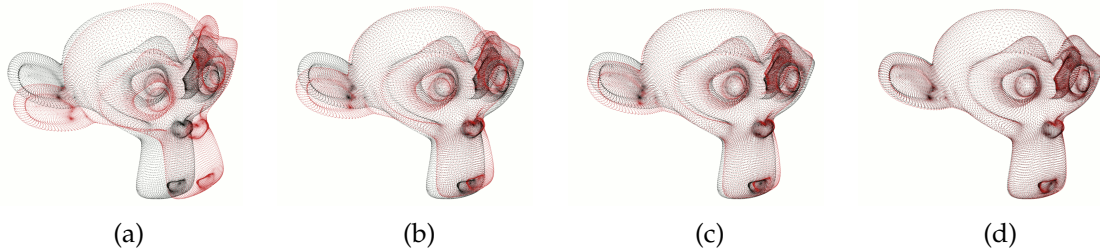

| (a) | (b) | (c) | (d) |

Figure 2.7: Iterative Closest Point (ICP) alignment after 0, 5, 10 and 20 iterations. The red point cloud converges to the black one with each iteration. The iterations are necessary since nearest neighbor searches are more likely to be able to find correct correspondences when the point clouds are more closely aligned [131].

2.3.4 Feature-based SLAM

To improve the performance of ICP, feature-based methods select specific points that are most likely to contain the most relevant information. After that, these points are also aligned through an iterative process such as ICP. Zhang and Singh in 2014, described an algorithm based on corner and plane points that are classified by the curvature of the surrounding points, to assist in fast classification and alignment [133]. In the following section, we will elaborate on their algorithm as it is a base for the work in this thesis. In a similar manner, Gelfand et al. compute the covariance matrix of each point in order to achieve geometrically stable matching [134]. Following this, Deschaud expands upon this concept and chooses those points that have a high degree of observability [135].

2.3.5 Surfel-based SLAM

The surfel-based map is a method of representing detailed geometric information of an underlying point cloud and is not dependent on particular feature formulations. As shown in Figure 2.9 (a), for each Surfel Element (surfel), a position, a normal, intensity (color), and a radius can be defined to form an ellipsoid and approximate a set of given points. Bosse and Zlot divide the volume into a grid of voxels and calculate a surfel for each of them [136]. Additionally, with an increasing number of grids, which are dynamically adjusted based on distance and robot position, they dynamically adjust the voxel size of the grid. In a similar way, Schadler et al. studied multi-resolution surfel maps, which are stored with an octree structure for fast access to the data [137] (see Figure 2.9 (b)). Alternatively, Behley and Stachniss save surfels as an unordered map along with the pose in order to make it easier to
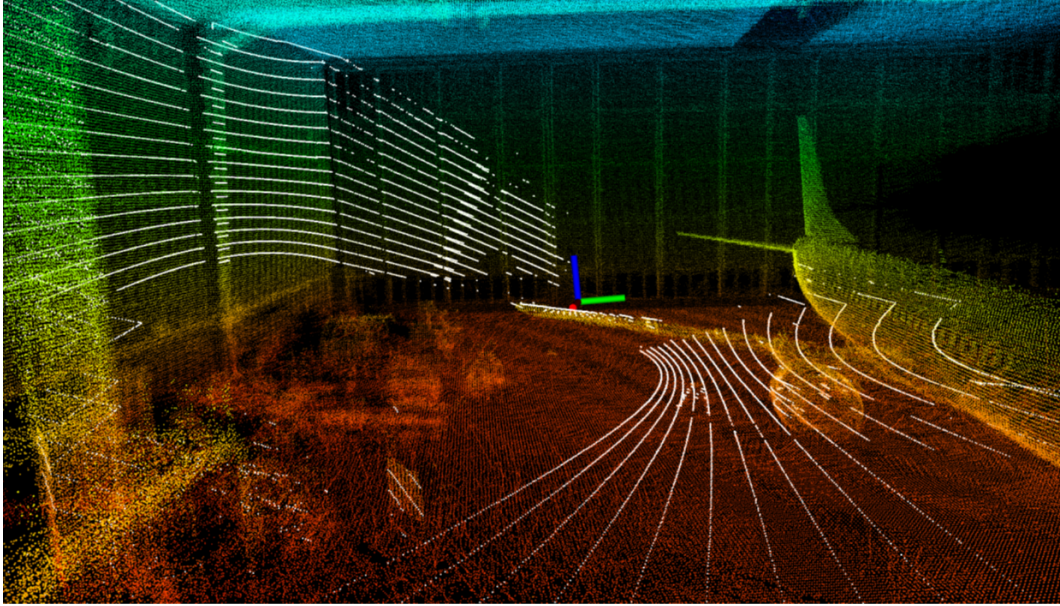
Figure 2.8: A sensor point cloud (white) is aligned to the map point cloud (colored) to determine the sensor's pose (three axis in red, green, and blue to show position and orientation). Height in the map is indicated with a gradient in color.

modify the position during loop closures [138]. Moreover, Park et al. have used what they call a dual surfel map, using both a 3D ellipsoid and a 2D disk surfel, [139].
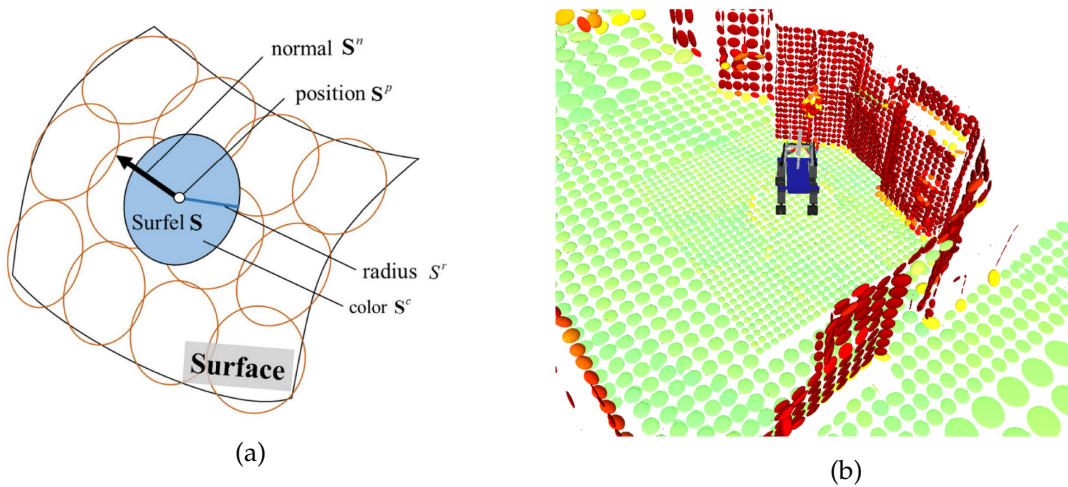

(a)   (b)

Figure 2.9: A multiresolution surfel map. Surfels closer to the robot have a higher resolution than those further away [137], [140]. (a) Schematic diagram showing the each attributes of a surfel. For each Surfel Element (surfel), a position, a normal, intensity (color), and a radius can be defined to form an ellipsoid and approximate a set of given points. (b) A Surfel Element (surfel) map with an octree structure [137].

2.3.6   LiDAR Odometry and Mapping (LOAM)

LiDAR Odometry and Mapping (LOAM) was first introduced by Zhang and Singh as a lightweight SLAM algorithm that can perform in real-time [141]. The approach is considered

state-of-the-art and leads the rankings for the KITTI autonomous driving benchmark [142] together with its spin-off Visual-LOAM (V-LOAM), which combines visual and laser ranging measurements [143]. The goal of this algorithm is to reduce the number of computations for simultaneous mapping and odometry by dividing the complex problem of SLAM into two sub-processes. The first algorithm performs odometry in real-time but with low fidelity to estimate the velocity of the LiDAR. The second algorithm runs at a lower frequency for map building and it returns a newly calculated part of the map that corresponds to the recent registered observations. Originally, pose estimation latency and update rate using LOAM is limited to the frequency of the sweeps of the laser scanner. However, many derivative researches have relied on the LOAM such as LEGO-LOAM [144] which is a lightweight and optimized version for ground navigation. Likewise, Yan et al. used LOAM to merge visual and laser data in their visual-LiDAR SLAM system [145]. While LOAM only implements an algorithm to minimize the odometry drift, full SLAM solutions including loop closure exist as well [146]. Further applications of LOAM exist for ground vehicles [147] [148], power line inspection [149], and search and rescue missions [150]. Subsequently, as a base of our work, we present the fundamental algorithm by Zhang [133].

2.3.6.1 System Overview

Several LiDAR systems can be utilized with LOAM, and the point clouds that result differ slightly. Single-line laser scanners rotate around one axis, but to get a full 3D view they must spin around another axis, perpendicular to the first. LiDAR systems can also be multi-lined with each laser scanner positioned at an individual angle. Those already produce a 3D point cloud without the need for additional actuation. However, the generated 3D point clouds using the multi-line LiDARs have limited vertical FOV, normally 30 to 45 degrees [151]. In the first approach, scan lines are crossed and recorded at differing times, whereas in the second approach, scan lines are parallel and recorded at the same time. In the following explanations, we refer to one full rotation of a laser scanner as a scan line and a set of related scan lines as a sweep. Real-time capability is achieved by splitting the algorithm into two parts. The odometry algorithm has a high frequency but low accuracy and is prone to drift. A mapping algorithm also runs at a lower frequency, but with higher accuracy and less drift. LOAM proposes frequencies of 10Hz and 1Hz respectively. The algorithms can be broken down into three main parts: scan registration, odometry, and mapping, which are explained in Figure 2.10. Since odometry and mapping are similar procedures, we will go into more detail about the feature point extraction algorithm below.

2.3.7 Feature Point Extraction

To find features within the observed point cloud, LOAM determines the points belonging to sharp edges and planar surface patches. To decide for a category, the smoothness c of each point $\mathcal{P}^L_{(k,i)}$ is calculated with the help of the set of adjacent points $\mathcal{S}$ within the same scan
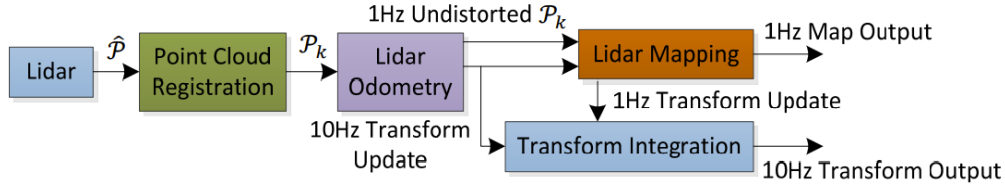
Figure 2.10: Overview of LiDAR Odometry and Mapping (LOAM). The LiDAR point cloud is first registered, then matched against the previous scan and later against the map. Based on the map the odometry pose is transformed for a high frequency output [45]. LOAM proposes frequencies of 10Hz and 1Hz respectively. The odometry algorithm has a high frequency but low accuracy and is prone to drift. A mapping algorithm runs at a lower frequency, but with higher accuracy and less drift.
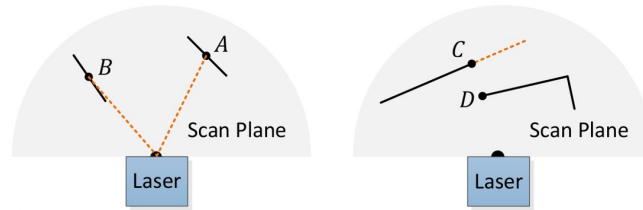
Figure 2.11: Feature Point Selection. Point A is selected as a feature point. Point B is discarded because of the angle to the surface (left). Furthermore, Point C and D are also omitted because they are part of an occluded scene (right) [133].

line by

$$c = \frac{1}{|\mathcal{S}| \cdot \left\| \mathcal{P}_{(k,i)}^L \right\|} \left\| \sum_{j \in \mathcal{S}, j \neq i} \left(\mathcal{P}_{(k,i)}^L - \mathcal{P}_{(k,j)}^L \right) \right\|, \tag{2.19}$$

where L is the LiDAR coordinate system and k is the sweep number. The c values are normalized so the distance to the laser scanner does not play a role in this context. All points are then categorized based on a threshold u, where points with a larger c value are edge points and points with a smaller c value are planar points. Furthermore, the flattest and sharpest points are sorted and selected based on the smoothness for the subsequent steps. To achieve a good distribution of those points, close neighbors to already selected ones are not considered. Additionally, each scan line is divided into multiple sub-regions, which can maximally contain two edge points and four planar points. In addition, points that are close to the boundary of an occlusion or that are oriented almost parallel to the laser beam are avoided [152].

2.3.8 Voxel Grid Map

In this section we investigate how new registered point clouds can be accumulated and builds up a map. The map is a voxel-filtered point cloud, which means the space of the cloud is divided into voxel elements, usually cubes, with a fixed size [153]. If there are multiple points inside one voxel, a centroid, the arithmetic mean position of the points, approx-

imates these points leaving only a single point in each voxel. This voxel-filtering effectively averages over the measurement errors, and thus the feature points are closer to the actual features they represent [154]. When a new point cloud set is added to the map, it is first filtered with the same voxel size as the map and then the points are added to the map point cloud [155]. This means some voxels now contain two or more points on average in the area the points were added. Filtering again induces that the new scan points and map point are valued 1:1 even though the map point represents the history of all the previous points in that voxel. Thus, a single bad registration can shift the map points aggressively. To reduce the possible map shift, the history of the previously observed points must be considered. Therefore, the ith map point $\mathcal{P}_{i,V}$ inside the voxel V can be calculated as

$$\mathcal{P}_{i,V} = \frac{(i-1)\mathcal{P}_{i-1,V} + S_V}{i}, \text{ with } \mathcal{P}_{1,V} = S_V \tag{2.20}$$

where S_V is the new scan point falling into the voxel V. This effectively reduces map shift and allows convergence towards the actual feature. Furthermore, the corresponding unique voxel key K_V from each point $S = (s_x, s_y, s_z)^T$ in the new scan is derived by

$$K_V = \begin{pmatrix} k_x \\ k_y \\ k_z \end{pmatrix} = \begin{pmatrix} \lfloor s_x \cdot r_x \rfloor \\ \lfloor s_y \cdot r_y \rfloor \\ \lfloor s_z \cdot r_z \rfloor \end{pmatrix}, \text{ with } k_x, k_y, k_z \in \mathbb{Z} \tag{2.21}$$

where $R = (r_x, r_y, r_z)^T$ is the resolution of the voxel grid given as the inverse of the voxel side lengths v_x, v_y and v_z, and the function $\kappa : K_V \to V$ is bijective. For fast access, we can save the voxel keys either in a red black tree or hash tables. While the red black tree preserves the geometrical structure, a hash-based access is generally faster. A hash key h is calculated quickly with the voxel key $K_V = (k_x, k_y, k_z)$ by

$$h = k_x + (k_y << s) + (k_z << 2s), \text{ with } h, s \in \mathbb{N}_0 \tag{2.22}$$

where $a << b$ is the bit shift operator, shifting the binary numbers a and b positions to the left and s to be the shift number, which is adjusted according to the voxel grid resolution for optimal performance. With this structure we only save the voxel key K_V, the latest map point $\mathcal{P}_{i,V}$, and the number of inserted points i^V. For each voxel at least one point was inserted in order to keeping the memory consumption at a minimum. Moreover, from the voxel grid an octree could be derived for fast nearest neighbor searches without the need to recompute a kd-tree after a new scan was added [156], [157].

2.4 Low-Latency 6D SLAM Using LiDAR-MIMU for Indoor Applications

Autonomous UAVs require a reliable navigation system to operate in challenging environments, such as GPS-denied or cluttered indoor areas [158], [159]. Tackling the general navigation mission, one needs to address a set of problems ranging from 6D pose estimation
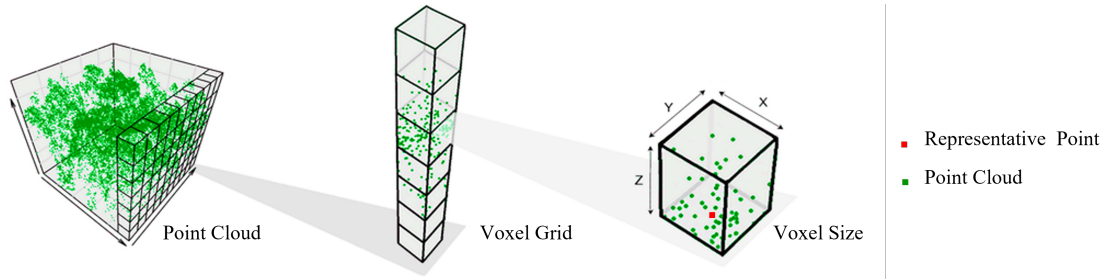
Figure 2.12: A voxel grid map is a voxel-filtered point cloud, which means the space of the cloud is divided into voxel elements, usually cubes, with a fixed size [153]. The values x,y, and z indicates the size of the voxel. If there are multiple points inside one voxel (green points), a centroid, the arithmetic mean position of the points, approximates these points leaving only a single point in each voxel (red point).
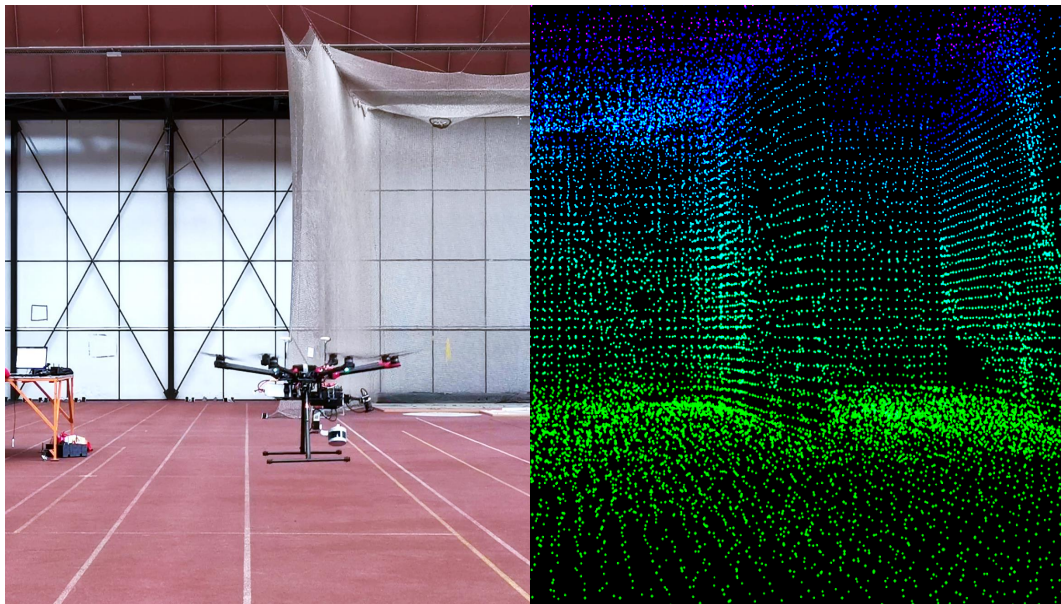

Figure 2.13: Real-time localization of a UAV in an indoor environment using the proposed Simultaneous Localization and Mapping (SLAM) framework. It provides low-latency 6D pose estimation with a high temporal update rate for autonomous navigation while generating a full-scale 3D map of the environment.

to trajectory planning [160], [161]. In practice, the navigation task's performance depends mainly on the accuracy, update rate, and latency of the perception unit. For instance, the trajectory tracking controller of a UAV requires a high-rate pose estimation (in the range of 0.1 to 1.0 kHz) with comparatively low latency (maximum up to 50 ms) to achieve a fully autonomous flight [162]–[164]. In this context, latency is the time difference from the sensor observation until the localization algorithm provides the estimated pose.

6D SLAM is mainly addressed in the literature by using multiple laser scanners, which are installed in both horizontal and vertical frames [141], [165]. However, due to the weight and processing constraints in UAV-based applications, in practice, the use of an actuated Li-DAR is preferred rather than installing multiple or heavier sensors with an inherently larger FOV [166], [167]. In this regard, different spinning mechanisms have been developed to be

mounted on drones [168], [169]. Although generally successful, existing solutions come with some shortcomings, such as long revisit time, blind spots, inconsistencies in scanning the environment, and skewing issues. These shortcomings cause problems in determining the correspondences within the point cloud data for 6D positioning [170], [171].

While LiDAR-based SLAM has been studied widely in the past decade for mobile robots [106], [141], [146], [170], [172], there is still a major gap in utilizing these systems for real-time navigation of UAVs. This is mainly because it is not possible to rely on an aerial vehicle remaining sufficiently motionless. The drone position must be represented at a high temporal update rate with low-latency to achieve stable closed-loop control. Addressing the challenge mentioned above, continuous-time trajectory estimators were developed in the literature [173], [174]. However, due to the limited FOV and the substantial time needed for a complete scan period of the LiDAR, estimating continuous odometry and global registration introduces a significant latency and low-fidelity 6D pose estimation in such a system. To address this issue, LiDAR-inertial SLAM systems use an additional Inertial Measurement Unit (IMU) to produce high-rate pose updates [106]. They provide acceptable results for ground-based mobile robots. However, as UAVs experience high dynamic motions and inevitable high-frequency vibrations due to the propellers' rotation, IMU-based approaches leads to having a fluctuating pose estimation in such applications [175].

To meet these challenges, within the scope of this work, we propose a low-latency localization and mapping framework using scan slicing [2]. In our approach, the point cloud data from a rotating laser scanner with 360° horizontal FOV is continuously sliced and used for estimating the real-time 6D pose. Besides, we introduce a novel sensor payload design in which a small, lightweight multi-line 360° laser scanner is actuated with a Lissajous pattern. This structure is employed to reduce the revisit time and provides a dense point cloud with consistency in scanning the environment [176], [177]. To this end, while UAV localization using actuated LiDARs has been presented previously [45], [173], [178], [179], this is the first work on scan slicing of an actuated rotating LiDAR, which is utilized for real-time low-latency indoor localization.

2.4.1   6D SLAM Using Actuated LiDAR

Commonly used multi-line LiDARs, like the ones from Velodyne or Ouster, usually have a vertical FOV between 15° and 45°. This perfectly fits the autonomous driving scenario because the car only moves in the xy plane and all relevant features are in sight. In literature, this is known as 3D SLAM. However, other systems like drones can move in 6D and rapidly change its motion. When mounting the laser scanner statically beneath the drone like one would do on a car, 6D pose estimation becomes a problem (see Figure 2.13). Because the LiDAR measures distances, enough laser beams have to hit the horizontal and vertical surfaces to achieve a sufficient accuracy. While in outdoor applications the downward measurement along with AHRS might be sufficient to estimate the 6D position. However, in an indoor environments, obstacles and walls can block the statically mounted laser scanner from registering points for vertical and horizontal surfaces. Additionally, for applications like the inspection of a large object (e.g, airplane), it is necessary to fly over the object, which would

greatly reduce the number of points recorded in the vertical direction. Considering a the previously covered issues for attitude and heading angle estimation using MIMUs and AHRSs in indoor environment, the 6D pose estimation becomes even more challenging for such applications. This implies a widespread FOV is more suitable for such operation, specifically considering the motion characteristic of mobile robots in general. Accordingly, Youji et al. state that the relevant features are based on the movement of the robot [147]. Therefore, as a base of this work, we rotate the laser scanner using external actuators perpendicular to the inner laser measurement angle to increase the FOV. This active actuation is common in drone-based applications with 2D range finders [44] [45], and yet concepts for typical passive actuation also exist [173]. Furthermore, manipulating the laser scanner solves another issue compared to the static mount: As long as the robot is stationary or relevantly motion-less the map consists solely of limited spaced out lines and only with applied motion the map can be extended. However, if the laser scanner is already actuated while the robot is encounters motion-less moments, a dense partial map of the environment is created, which helps to detect the velocity and the 6D pose of the robot with more confidence.

2.4.1.1   Actuation Models

We determined several models to actuate the laser scanner. The typical one is a rotation around a single axis where one can choose different axes. We modeled two actuations within the scope of this work. The first one is a $\pm 40°$ rotation around the vertical axis. The other one is a $\pm 40°$ rotation around both horizontal and vertical axes. These methods increases the vertical FOV from 30° to 120°. Although Yoshida et al. developed a motion that can be controlled by a single motor, it can also be performed with rotations around the pitch and roll axes [180] independently. Because we perform the movement with two motors, the additional degree of freedom allows us to dynamically change the horizontal and vertical observation time. We calculate the angles for the motors at every time point t with

$$pitch(t) = \alpha(t) \cdot \sin(\psi(t))$$
$$roll(t) = \alpha(t) \cdot \cos(\psi(t))$$

(2.23)

where $\alpha(t)$ is the opening angle and $\psi(t)$ the rotational angle. Furthermore, as is shown in Figure 2.14, the two-axes actuation creates a more uniform distribution of measurements than the single axis one as previously discussed.

2.4.2   MIMU Utilization

LiDAR-only pose estimation process underestimates acceleration and the angular velocity for 6D ego-motion estimation [181]. This is mainly is because of data acquired during continuous vehicle motion suffer from geometric motion distortion and can subsequently result in poor metric estimates. An inertial measurement unit (IMU) in turn is equipped with multi-sensors and the fusion of MIMU and other sensor data has been proven to increase localization performance [103], [106]. Even though the original LOAM utilized an acceleration from an IMU, attitude and heading estimation from multi-MIMU support is not included,
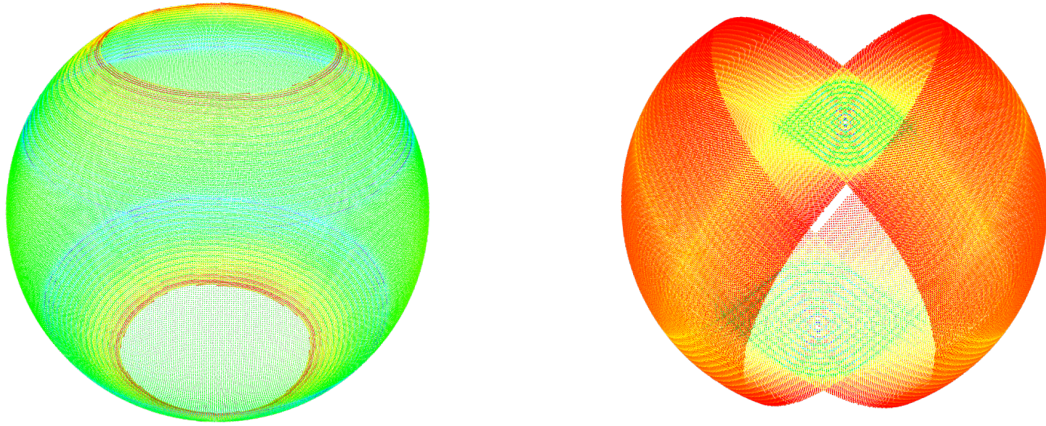
Figure 2.14: Distribution of measurements on the unit sphere based on the actuation. Point clouds are colored based on the density of measurements. A sinusoidal movement around one axis has a high density close to the axis (green) and a lower one outside (red) (left). A rotation based on Equation 2.23 is evenly distributed (right). Both point clouds have an α of 40° [2].

and thus we establish how to integrate it into the SLAM structure. To get a reliable attitude, heading, and the position from the multi-MIMU measurement one has to fuse the data using probabilistic approaches, which means errors accumulate quickly. However, this can be compensated with the LiDAR data. While the strength of the ego-motion estimation process lies in the linear movement estimation, the strength of the multi-MIMU is the detection of changes in angular and translational movements. We therefore combine the multi-MIMU measurements and the 6D ego-motion estimation with an EKF and set the result as a prior for the scan to map matching process. This combined prior improves the mapping as a better prior enables the convergence to the right local minimum. Furthermore, a prior closer to the local minimum also reduces the optimization time as fewer iterations are needed because the corresponding points from the two point sets are already closer to each other.

2.4.3 Low-Latency Ego-Motion Estimation

State-of-the-art laser scanners rotates with an average of 1200 Rounds Per Minute (RPM) and provides up to 1,000,000 points per second. In a best possible way, a 360° scan is available every $\omega_{scan} = 50$ milliseconds and contains several thousand measurements. The laser scanner is continuously measuring and creates a constant output stream associated with time and angle. These measurements are accumulated over time at the driver level. When the laser scanner completes a full 360° scan, it releases all the recorded points as a single sweep. State-of-the-art SLAM frameworks are developed based on these 360° scans for ego-motion estimation[146], [172], [175]. In practice, the point cloud arrival rate defines the frequency of the pose estimation. However, the ego-motion estimation can be performed after any other measurement, which means 360° is an arbitrary margin, and individual measurements can be bundled freely.

Although it is reasonable to compare complete 360° scans in the odometry unit; we show

later in the proposed localization framework that our SLAM model does not depend on the odometry unit in comparison to the typical models. The proposed SLAM framework uses continuous slices of a full 360° scan for low-latency localization using direct scan matching on the global map. The main intention of the proposed Continuously Slicing Point Cloud (CSPC) is to gather smaller collections of measurements that can be used in a successive manner instead of waiting for a full 360° scan. In this context, describing the term continuous in other words, we collect the measurements from the upcoming scan into a small point cloud slice for the last θ angle (i.e., only the data of the measurements within the last $\theta = 45°$ of the rotation), and provide them one after another to the ego-motion estimation process. Slicing of the points is developed in the driver level of the laser scanner, and the CSPC is published at a rate of $\frac{1}{\Delta q}$ Hz. We utilized message publishing system developed by Robot Operating System (ROS). Considering each slice is generated from the last $\theta = 45°$ of the rotation of the LiDAR, each slice is available every $\Delta q = 6.25$ ms and noted by $\mathcal{P}_q$ with q indicating the last observed slice. With less angular coverage, the point cloud becomes more ambiguous because opposing features are missing to prevent misguided shifts during optimization in the pose estimation. However, the new measurements can be integrated with old ones to achieve the necessary number of measurements to ensure correct matching. For instance, we can add the new slice ($\theta = 45°$) to the previously registered slices within the last 315°, in order to obtain a repeatedly full 360°scan, however, at a much higher update rate. While the reuse of past measurements does not affect the achieved update rate, the system's real-time performance needs to be carefully investigated due to the increase in computational usage.

2.5   Chapter Summary

This chapter introduced the most relevant background needed to better comprehend the reasoning discussed and the concepts proposed in this work. The first part addressed the basics of sensor systems, which are then used to detail the acquisition strategies for the creation of multi-MIMU fusion and point cloud based SLAM. The transition from external reference systems to laser scanning data acquisition is made to emphasize the challenges of acquiring LiDAR-MIMU footage. State-of-the-art technologies that approach the attitude and heading reference estimation are presented and thoroughly debated. Further topics that thematically relate to the overall goal of low-latency 6D SLAM, such as Feature-based SLAM, are further introduced and discussed. The prior arts' advantages and drawbacks are compared to the properties of the proposed low-latency 6D SLAM using LiDAR-MIMU system.

Chapter 3

Skewed-redundant Hall-effect Magnetic Sensor Fusion for Perturbation-free Indoor Heading Estimation

3.1 Perturbation-free Heading Reference Estimation based on Skewed Redundant MIMU Fusion

It is an essential component of indoor localization in robotic applications to accurately estimate the current attitude and heading in relation to a known reference. Affordable Attitude and Heading Reference System (AHRS), typically use 9-axis solid-state MEMS-based sensors to determine attitude and heading. Such a system relies on the accuracy of measuring the Earth's magnetic field to estimate heading. In contrast, measures of the Earth's magnetic field using MEMS-based magnetometer sensors in an indoor environment are strongly affected by external magnetic disturbances. In this chapter, we present a novel approach for robust indoor heading estimation using skewed redundant magnetometer fusion. Based on hall-effect magnetic sensors, a tetrahedron-shaped platform is designed for the purpose of determining Earth's magnetic field and for compensating for anomalies in external magnetic fields. In addition, a correlation-based fusion technique is presented to mitigate perturbations using the proposed skewed-redundant configuration. In the proposed fusion technique, we use a correlation coefficient analysis to determine the distorted axis and extract the perturbation-free vector of Earth's magnetic field from the redundant magnetic measurements. According to our experimental results, the proposed scheme is capable of effectively mitigating the anomalies in magnetic field measurements, and is capable of accurately estimating the Earth's true magnetic field. This evaluation was the base for the deep-fusion based approach where covers a more comprehensive solution to attitude and heading estimation covered in the consecutive chapter.

3.2 Authors Contribution

Development of the measurement method using tetrahedron-shaped platform, a correlation-based fusion technique, and execution of the experiments were performed by Mojtaba Karimi. Data visualization were conducted by Edwin Babaians. The printed circuit board was conceptualized by Mojtaba Karimi. Martin Oelsch, Tamay Aykut, and Mojtaba Karimi evaluated the experimental results. Eckehard Steinbach supervised this work. The original draft was written by Mojtaba Karimi. The paper was reviewed and edited by all authors.

Title:

Skewed-redundant Hall-effect Magnetic Sensor Fusion for Perturbation-free Indoor Heading Estimation

Authors:

Mojtaba Karimi, Edwin Babaians, Martin Oelsch, Tamay Aykut, and Eckehard Steinbach

Skewed-redundant Hall-effect Magnetic Sensor Fusion for Perturbation-free Indoor Heading Estimation

Mojtaba Karimi[1], Edwin Babaians[1], Martin Oelsch[1], Tamay Aykut[2], and Eckehard Steinbach[1]

[1]Department of Electrical and Computer Engineering, Technical University of Munich
Munich, Germany

[2]Department of Electrical Engineering, Stanford University
Stanford, CA, USA

Email:[1]{mojtaba.karimi, edwin.babaians, martin.oelsch, eckehard.steinbach}@tum.de, [2]{tamaykut}@stanford.edu

Abstract—Robust attitude and heading estimation with respect to a known reference is an essential component for indoor localization in robotic applications. Affordable Attitude and Heading Reference Systems (AHRS) are typically using 9-axis solid-state MEMS-based sensors. The accuracy of heading estimation on such a system depends on the Earth's magnetic field measurement accuracy. The measurement of the Earth's magnetic field using MEMS-based magnetometer sensors in an indoor environment, however, is strongly affected by external magnetic perturbations. This paper presents a novel approach for robust indoor heading estimation based on skewed-redundant magnetometer fusion. A tetrahedron platform based on Hall-effect magnetic sensors is designed to determine the Earth's magnetic field with the ability to compensate for external magnetic field anomalies. Additionally, a correlation-based fusion technique is introduced for perturbation mitigation using the proposed skewed-redundant configuration. The proposed fusion technique uses a correlation coefficient analysis for determining the distorted axis and extracts the perturbation-free Earth's magnetic field vector from the redundant magnetic measurement. Our experimental results show that the proposed scheme is able to successfully mitigate the anomalies in the magnetic field measurement and estimates the Earth's true magnetic field. Using the proposed platform, we achieve a Root Mean Square Error of 12.74° for indoor heading estimation without using an additional gyroscope.

I. INTRODUCTION

Attitude and heading estimation is one of the fundamental requirements for robotics, human machine interaction, and navigation in indoor environments [1]–[3]. The ability of self-localization for autonomous systems (e.g., unmanned aerial vehicles, mobile robots) or prediction of head motion for teleportation or virtual reality applications [4], [5], similarly, crucially depend on the reliable estimation of attitude and heading angles [6]–[8]. The result of camera-based or LiDAR-based SLAM systems, for instance, becomes more reliable by utilizing robust attitude and heading information which can be attained mainly from an Attitude and Heading Reference System (AHRS) or an Inertial Measurement Unit (IMU) [9]–[11]. The main difference between an AHRS in contrast to the IMU is the preprocessing of the sensory data in order to provide absolute orientation estimation with respect to the Earth's geometrical characteristics. The known reference of AHRS comes from the Earth's gravitational and magnetic field. In this regard, attitude estimation refers to the angular measurement on the vertical plane, with respect to the local
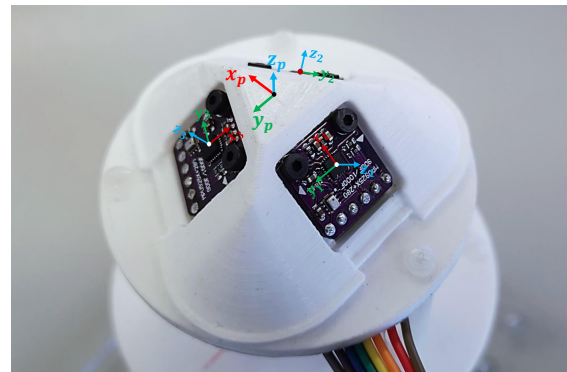

Fig. 1: Skewed-redundant magnetometer platform. The measurement point p is defined as the virtual center of the multi-magnetometer measurement system. A correlation-based fusion method is used for the mitigation of external magnetic field anomalies and robust heading estimation in an indoor environment.

level frame, and is computed as inclinations. Similarly, angular measurement on the horizontal plane, with respect to the geographical true north, is referred as heading angle (also known as yaw or azimuth) [12].

Affordable AHRSs are typically using small-scale sensors based on Micro-Electro-Mechanical Systems (MEMS). The MEMS-based sensors have the advantages of low cost, low power consumption, and high update rate. A general 9-DoF AHRS consists of a 3-DoF accelerometer, a 3-DoF gyroscope, and a 3-DoF magnetic field meter. Exploiting the complementary nature of the MEMS-based inertial sensors, 3-DoF accelerometers along with 3-DoF gyroscopes are mainly fused for attitude estimation [13]. Similarly, the MEMS-based AHRSs are mostly using a single 3-DoF magnetometer to measure the magnetic field strength and utilizing the previously determined relative calibration values to extract the Earth's geographic azimuth. In practice, the azimuth estimation from the magnetometer is fused with the gyroscopes to provide more stable heading estimation. Nonetheless, the general accuracy for heading estimation on such a AHRS highly depends on the Earth's true magnetic field measurement.

Determining error-free and reliable heading estimation with

respect to a known reference is problematic in case of indoor applications. This is mainly because of different sources of errors in the MEMS-based magnetometer measurements [14]–[17]. The most dominant source of error for such a magnetometer sensor, however, is the interference of the external magnetic field in an indoor environment. Permanent magnets and ferromagnetic materials are the main cause of such a magnetic perturbation. In this regard, calibration of the MEMS-based magnetometers for heading estimation has been investigated thoroughly during the last decades [14], [18]. Probabilistic models for sensor calibration and fusion such as the Extended Kalman Filter (EKF), the Unscented Kalman Filter (UKF), and Ellipsoid Fitting are well known in this field [19]–[22]. However, dynamic magnetic deviation caused by external anomalies cannot be fully compensated using the sensor calibration methods. To overcome the aforementioned problem, redundant sensor systems have shown promising results as they are using redundant measurements to determine the source of error and increase the estimation accuracy [23]–[25].

Multiple magnetometers in combination with gyroscopes were used to approximate the external magnetic disturbances and to provide perturbation-free heading estimation [24], [26]–[28]. In these works, the authors used an estimation of the expected magnetic field using the temporary rotation vector to compensate the external magnetic perturbation. These approaches need a precise and simultaneous calibration of the magnetometers and the gyroscopes to be able to provide an acceptable heading estimation. In [29], a magnetic heading compensation method for disturbances introduced by near-surface interference was proposed. The authors used inverse magnetic anomaly vector estimation to compensate for the external perturbation. The main limitations of this approach are the predefined magnetic field model and one-dimensional perturbation mitigation. In general, these approaches are usually bulky, expensive, and often rely on expensive calibration procedures. Nonetheless, it has been shown that the accuracy of the magnetic field measurement is significantly improved by using redundant magnetometer sensors [30]–[33]. However, the non-deterministic errors cannot be completely removed in real application scenarios due to their transient nature. This is mainly because of limitations in the system identification procedure, non-linear dynamic errors, and the use of strict calibration models. The skewed-redundant inertial sensor fusion, on the other hand, provides more reliable output, as shown in [34]–[37]. The skewed axes configuration for redundant sensor geometry has been widely investigated for attitude estimation by applying this method on accelerometers and gyroscopes. However, this method has been poorly investigated for the redundant magnetometers.

Inspired by the success of redundant integrated sensors for processing high-dimensional sensor data [24], [26], [38], and the benefits of redundant, non-orthogonal sensor structures [38]–[41], we propose a skewed-redundant magnetometer platform for perturbation-free heading estimation. The approach undertaken in this paper relies on the use of redundant sensor
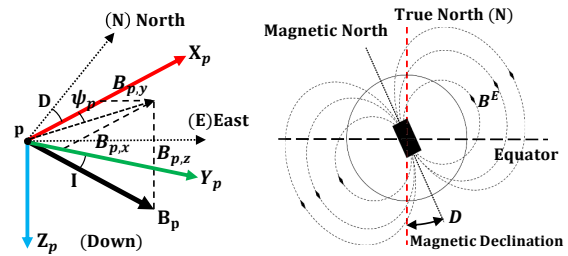

Fig. 2: Components of the Earth's magnetic field and the abstraction of the heading estimation using a 3-axis magnetometer.

information to enhance the global performance of the heading estimation while ensuring the system's integrity. Unique non-orthogonal placement of the triple 3-axis magnetometer is designed based on the Hall-effect sensors principle in order to mitigate the external magnetic perturbation by extracting the distracted signals. The proposed approach uses a correlation-based sensor fusion method for external magnetic perturbation mitigation and robust heading estimation (Fig. 1). The skewed-redundant magnetometer platform and correlation-based filter and fusion strategy shows substantial improvement in heading estimation accuracy in an indoor environment. Our contributions can be summarized as follows:

- We introduce a skewed-redundant magnetometer platform based on the Hall-effect magnetic sensor principle in order to mitigate the external magnetic perturbation.
- We propose a filter model based on correlation analysis and investigate the proposed platform for reliable estimation of the Earth's magnetic field.
- Moreover, we validate our approach by means of using the proposed setup in an indoor environment to prove its general validity. We show its superior performance in contrast to the naive approach or KF-based fusion, where no perturbation compensation is applied.

II. MAGNETIC FIELD COMPONENTS AND HEADING ESTIMATION USING MAGNETOMETERS

A combination of a 3-axis accelerometer, a 3-axis gyroscope, and a 3-axis magnetometer is typically used in low-cost AHRS sensors. The estimated angles are defined as the corresponding Euler angles roll (α), pitch (β), and yaw (γ). In order to improve the robustness of the roll and pitch estimation, the measurements provided by the gyroscope and accelerometer are typically fused. The magnetometer, on the other hand, can be complemented using the gyroscope for heading estimation. Heading estimation with respect to a known reference can be performed by considering the fact that solid state magnetometers are used for measuring changes in the strength of the Earth's magnetic field. Dominant technologies for the solid state MEMS-based magnetometers are based on the Hall-effect or the Magneto Resisting effect [17], [42]. Considering that the magnetic field is a vector field, it is characterized by its magnitude and direction or its components

along the coordinate axes. A sensor with at least three axes is required to determine the Earth's magnetic field vector (Fig. 2). In a Cartesian coordinate system at the measurement point p, the magnetic field vector $\mathbf{B}_p$ can be defined as

$$\mathbf{B}_p = B_{p,x}\,\boldsymbol{i} + B_{p,y}\,\boldsymbol{j} + B_{p,z}\,\boldsymbol{k} \;, \tag{1}$$

where the x-axis is oriented along the geographical meridian, and the direction to the north is positive, the y-axis along the parallel with positive direction toward the east and the z-axis is directed downward. The observation point p is the origin of the measurement coordinate system. The angle I between the horizontal plane and the vector $\mathbf{B}_p$ is called the inclination [43]. The local magnetic declination D, sometimes called magnetic variation, is the angle between the magnetic north and the true north. Accordingly, the Earth's geographic azimuth ψ_p can be derived by considering the local declination D and the perpendicular components of the measurement vector $\mathbf{B}_p$ as

$$\psi_p = \tan^{-1}\left(\frac{B_{p,y}}{B_{p,x}}\right) \; \pm D \;. \tag{2}$$

The magnetic field of a dipole, similarly, is characterized by its magnetic moment $\mathbf{M}$. The strength and direction of this field depends not only on the magnitude of the magnetic moment, but also on its orientation relative to the direction of the magnetic field [43]. In the presence of multiple dipoles, the magnetic field $\mathbf{M}_g$ is defined by the sum of all magnetic moments $\mathbf{M}_i$ and is given by

$$\mathbf{M}_g = \sum \mathbf{M}_i \;. \tag{3}$$

The curl of the magnetic vector potential, the magnetic field $\mathbf{B}_p$, is governed by the density of magnetic moments at the measurement point and is derived classically using the Biot-Savart law [44]. The combined field can change the measurement field vector $\mathbf{B}_p$ due to the presence of multiple dipoles. This change in the magnetic field is the leading cause of heading errors in an indoor environment. As shown in Fig. 3, the effect of the external magnetic field on the p-type Hall element depends highly on the orientation and the distance of the external magnetic moment to the perpendicular sensing axis [45]. Moreover, a minimum number of three Hall elements perpendicular to each other is required to determine the magnetic field $\mathbf{B}_p$ as it is shown in Eq. 1. This implies that the measuring axis with the lowest angle to the external magnetic anomaly will report a non-correlated measurement compared to the other axes. Additionally, in a perturbation-free environment, the norm of the magnetic field vector measurement should be equal to the magnitude of the Earth's magnetic field. This norm, H_p can be extracted from the Earth's geomagnetic model [46]. The aforementioned principles are the main motivation of proposing the skewed-redundant magnetometer platform (see Fig. 1).
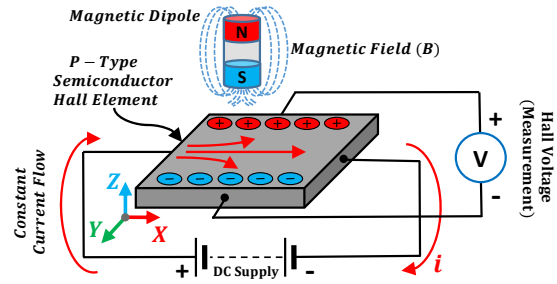

Fig. 3: Single-axis Hall-effect sensor principle [45]. The output signal from a Hall-effect sensor is a function of the magnetic field density around the device. A minimum of three perpendicular axes is needed for the magnetic field vector measurement.
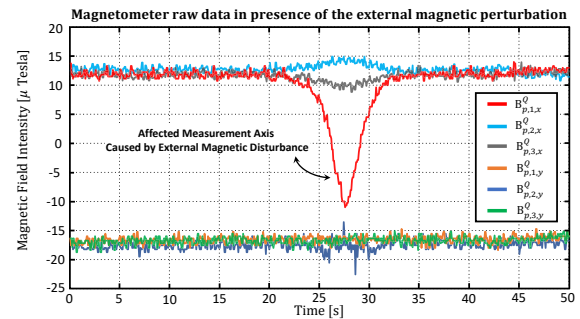

Fig. 4: Output signals from a skewed-redundant magnetometer sensor at measurement point p and the effect of the external disturbance on the specific sensing axes, which are perpendicular or have a small angle to the source of the external magnetic moment.

III. Skewed-redundant Magnetometer Platform

A significant outcome in [13] is that an ideal setup of redundant inertial sensors in terms of fault tolerance is a "skewed" configuration in which the delicate axes of each tri-axial sensor are not aligned and transformations are performed to align them in a virtual coordinate frame. This method has been examined on inertial sensors such as accelerometers as well as gyroscopes and showed a significant improvement in the measurement accuracy [34], [36], [37]. Following the concept of the skewed-redundant inertial sensor, we propose a skewed-redundant magnetometer platform and analyse its performance.

From the sensor data, it is observable that the perpendicular axis to the external magnetic anomalies can be affected dramatically. The pattern of this change in the data is recognizable, and this motivates us to examine different mounting structures with redundant magnetometers. It is difficult to determine which configuration yields the best performance since the large number of possible arrangements makes it difficult to derive a deterministic relation between the system errors and the configuration parameters. We investigated more
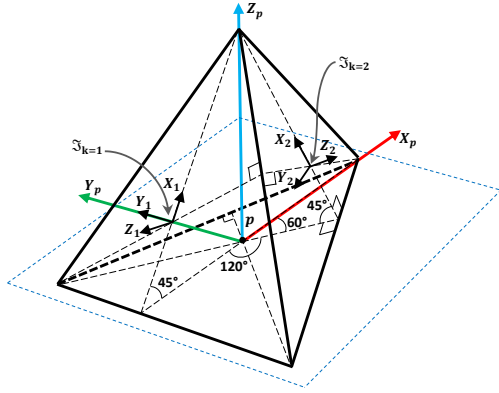
Fig. 5: Tetrahedron configuration model of the skewed-redundant magnetometer platform using three 3-axis magnetometer sensors. The virtual coordinate frame for the measurement are shown with the center point p, and only two sensor frames are depicted for simplification.

TABLE I: Formation of the extracted Direction Cosine Matrix for the skewed-redundant tetrahedron platform depicted in Fig. 5.

Frame	Direction Cosine Matrix (DCM)
$R^p_{\Im,1} =$	$\begin{bmatrix} -cos(45°) & 0 & sin(45°) \\ 0 & -1 & 0 \\ sin(45°) & 0 & cos(45°) \end{bmatrix}$
$R^p_{\Im,2} =$	$\begin{bmatrix} cos(45°)cos(-60°) & -cos(45°)sin(-60°) & sin(45°) \\ sin(-60°) & cos(-60°) & 0 \\ -sin(45°)cos(-60°) & sin(45°)sin(-60°) & cos(45°) \end{bmatrix}$
$R^p_{\Im,3} =$	$\begin{bmatrix} cos(45°)cos(60°) & -cos(45°)sin(60°) & sin(45°) \\ sin(60°) & cos(60°) & 0 \\ -sin(45°)cos(60°) & sin(45°)sin(60°) & cos(45°) \end{bmatrix}$

than ten different configurations and analyzed the data to determine for which of them the external disturbance is perceivable by considering the minimum number of axes and the coverage orientation. The skewed configuration in comparison to the orthogonal structures showed better results. Using a tetrahedron skewed configuration [47], we recorded the data shown in Fig. 4, which illustrates that the axis which has the minimum angle difference to the source of the external magnetic perturbation, reports uncorrelated output while the other axes of the magnetic measurements are less impacted.

To determine the Earth's true magnetic field, we consider the measured magnetic field as a combination of the external magnetic perturbation and the Earth's magnetic field. Taking Eq. 3 into consideration, the measured magnetic field $\mathbf{B}^Q$ in the sensor frame with center point $\Im$ can be defined as

$$\mathbf{B}^Q_\Im = \mathbf{B}^E_\Im + \mathbf{B}^\varepsilon_\Im \ , \tag{4}$$

where $\mathbf{B}^\varepsilon_\Im$ is the external magnetic field disturbance and $\mathbf{B}^E_\Im$ is the local Earth's magnetic field. To have all the true signals in the measurement point p, sensor data needs to be aligned

to a defined virtual coordinate frame with the center point p. Additionally, we normalize the measured magnetic field vector to be able to use the correlation analysis later for sensor filtering and fusion. We can define $\mathbf{B}^Q_{p,k}$ for each sensor k using

$$\mathbf{B}^Q_{p,k} = \mathbf{R}^p_{\Im,k} \left[\frac{B^Q_{\Im,k,x}}{\left\| \mathbf{B}^Q_{\Im,k} \right\|} \quad \frac{B^Q_{\Im,k,y}}{\left\| \mathbf{B}^Q_{\Im,k} \right\|} \quad \frac{B^Q_{\Im,k,z}}{\left\| \mathbf{B}^Q_{\Im,k} \right\|} \right]^T , \tag{5}$$

where $\mathbf{R}^p_{\Im,k}$ is the Direction Cosine Matrix (DCM) with respect to roll (α), pitch (β), and yaw (γ) angles. The angles roll, pitch, and yaw are the representative rotation of the 3-axis magnetometer from the sensor coordinate frame with the center point $\Im$ around the virtual coordinate frame with the center point p. Based on the skewed configuration, the proposed structure with the mounting angles is depicted in Fig. 5. A rotation matrix can be defined for an individual sensor k according to the static angles using

$$\mathbf{R}^p_{\Im,k} = \begin{pmatrix} c\beta c\gamma & -c\beta s\gamma & s\gamma \\ c\alpha s\gamma + s\alpha s\beta s\gamma & c\alpha c\gamma - s\alpha s\beta s\gamma & -s\alpha c\beta \\ s\alpha s\gamma - c\alpha s\beta c\gamma & s\alpha c\gamma + c\alpha s\beta s\gamma & c\alpha c\beta \end{pmatrix} . \tag{6}$$

Similarly, Table I shows the extracted Direction Cosine Matrix from the generic rotation matrix defined in Eq. 6 for each of the three sensors in the tetrahedron skewed-redundant platform.

IV. SKEWED-REDUNDANT MAGNETOMETER FUSION

The main source of magnetometer errors are the sensor offsets, scale factor, non-orthogonality, and magnetic deviation [24], [48], [49]. The latter is based on the magnetization characteristics of the sensor and can have a static or dynamic source. The static magnetic deviation error consists of soft- and hard-iron parts. In this paper, we are not going to the details of the sensor hard-iron and soft-iron calibration procedure. We assume the measured magnetic field $\mathbf{B}^Q_{p,k}$ is calibrated by using the method described in [14]. Additional to the static magnetic deviation error which can be addressed by applying the offline calibration, the dynamic magnetic deviation error is still problematic and subsists as the most dominant error source. The proposed multi-magnetometer fusion, therefore, uses dynamic weighting based on the normalized correlation coefficient analysis to filter the axes affected by the external disturbance. This mitigates the external dynamic magnetic field perturbation and allows for the estimation of the true Earth's magnetic field. Let us define the true Earth's magnetic field vector $\mathbf{B}^F_p$ as

$$\mathbf{B}^F_p = \sum_{k=1}^n \mathbf{J}_k \mathbf{M}_k (\mathbf{S}_{p,k} \mathbf{B}^Q_{p,k} + \mathbf{b}_{p,k} + \mathbf{W}_{p,k}(t)) \ , \tag{7}$$

where $n = 3$ is the number of sensed magnetic field vectors that are transformed to the measurement point p in the skewed-redundant multi-magnetometer setup, $\mathbf{S}_{p,k}$ is the scale factor of the sensor k, $\mathbf{b}_{p,k}$ is the bias vector, $\mathbf{W}_p(t)$ is the white noise vector, $\mathbf{M}_k$ represent the skew-symmetric misalignment matrix and $\mathbf{J}_k$ is the adaptive correlation-based identity dynamic weighting matrix.

We use the general definition of the $\mathbf{S}_{p,k}$ and $\mathbf{M}_k$ for the proposed magnetometer sensor setup from [31], and [25]. It should be noted that separating all components of the encountered errors for the magnetometer platform requires unique setups (e.g., turning tables) and is hard to accomplish. However, this is not essential in the case of the MEMS-based sensors because some error elements predominate and the error model can, therefore, be simplified [50], [51].

The biases $\mathbf{b}_{p,k}$ are approximated for each 3-axis magnetometer sensor and can be defined as

$$\mathbf{b}_{p,k} = \mathbf{b}_k^{\mathrm{C}} + \mathbf{b}_k^{\mathrm{RC}} + \mathbf{b}_k^{\mathrm{GM}}(t) + \mathbf{b}_k^{\mathrm{RW}}(t) + \mathbf{b}_k^{\mathrm{BI}}(t) \ , \qquad (8)$$

where $\mathbf{b}_k^{\mathrm{C}}$ are the constant components of the bias vector and can be extracted from sensor design characteristics, $\mathbf{b}_k^{\mathrm{RC}}$ is the stochastic process with maximum constant value which is extracted from the hard-iron calibration procedure and having the distribution $\mathbf{b}_k^{\mathrm{RC}} \backsim N(\mu_k^{\mathrm{b}}, \sigma_k^{\mathrm{b}2})$. The $\mathbf{b}_k^{\mathrm{GM}}$ is a first-order Gauss-Markov stochastic process applied on the sensor constant bias estimation. Constant biases of the sensors are extracted from the standard deviation of the measurements considering that each measurement is a discrete sample measurement at t [52]. Therefore, $\mathbf{b}_k^{\mathrm{GM}}$ is defined as

$$\dot{\mathbf{b}}_k^{\mathrm{GM}}(t) = \zeta . \mathbf{b}_k^{\mathrm{GM}}(t) + \mathbf{E}_k(t) \ , \qquad (9)$$

where $\zeta = 1/T$ is the inverse of the correlation time with the initial condition

$$\lim_{t_0 \to \infty} \mathbf{b}_k^{\mathrm{GM}}(t_0) = 0 \ ,$$

and $\mathbf{E}_k(t)$ is a zero-mean Gaussian white-noise process which is applied on the sensor constant bias vector. We denote the finite bandwidth of the respective magnetometer data by Δf, which can be calculated from the auto-correlation function as expressed in [53]. The Nyquist-Shannon sampling rate equals $2\Delta f$ which is equal to two samples per period T. Hence, defining $T/2$ as the correlation time, the Nyquist sampling rate ζ is equal to the inverse of the correlation time. The $\mathbf{b}_k^{\mathrm{RW}}(t)$ is a random walk stochastic process defined as

$$\dot{\mathbf{b}}_k^{\mathrm{RW}}(t) = \mathbf{E}_k(t)$$

with the initial condition $\mathbf{b}_k^{\mathrm{RW}}(t_0) = 0$. The $\mathbf{b}_k^{\mathrm{BI}}(t)$ are the bias instabilities which were added to the original model developed in [30] and [54] to describe more precisely the noise characteristics observed for the skewed-redundant magnetometer platform (see Fig. 5). We use the same definition of the proposed bias instabilities vector as in [30]. The white noise process $\mathbf{W}_{p,k}$ is defined as

$$\mathbf{W}_{p,k}(t) = \sigma_{p,k}^{\mathrm{w}} \mathbf{E}_k(t) \ , \qquad (10)$$

where $\sigma_{w,k}^{\mathrm{w}}$ are the standard deviation of the measurements. The fixed parameters (i.e., the constant components of the scale factors) and bias vector components were estimated with a classical least-squares parametric compensation. The remaining parameters were estimated with variance analysis techniques, as explained in [30] and [55].
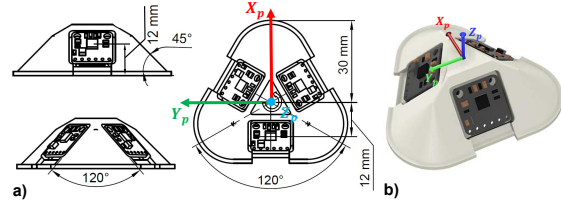

Fig. 6: Structure of the 3D printed model of the tetrahedron platform with the illustration of the dimension of the module.

The adaptive correlation-based identity dynamic weighting matrix $\mathbf{J}_k$, which is used for identifying the external perturbation using the correlation analysis is defined as

$$\mathbf{J}_k = \begin{pmatrix} \frac{\omega_{k,x}}{\|\omega_{k,\mathbf{X}}\|} & 0 & 0 \\ 0 & \frac{\omega_{k,y}}{\|\omega_{k,\mathbf{Y}}\|} & 0 \\ 0 & 0 & \frac{\omega_{k,z}}{\|\omega_{k,\mathbf{Z}}\|} \end{pmatrix}, \qquad (11)$$

while the individual component of the dynamic weight vector ω_k is defined as the correlation between each axis and the corresponding axes measuring the same direction of the magnetic field in the virtual measurement point p. For the triple measurement, we can define

$$\omega_k = \left[\left| \rho_{x_\eta, x_a x_b} \right| \quad \left| \rho_{y_\eta, y_a y_b} \right| \quad \left| \rho_{z_\eta, z_a z_b} \right| \right]^T, \qquad (12)$$

while the normalization term in the identity weighting matrix for the number of n sensors can be defined from all the corresponding axes using

$$\|\omega_{k,\mathbf{N}}\| = \sum_{i=1}^{n} |\omega_{i,\mathbf{N}}| \ . \qquad (13)$$

Eventually, the joint correlation coefficient ρ between the current axis measurement η and the two similar non-orthogonal axis a and b, which are rotated using the $R_{\Im,k}^p$ from the measurement sensor frame with the center point $\Im$ to a virtual frame with the center point p, can be calculated by

$$\rho_{\eta, ab} = \sqrt{\frac{r_{\eta,a}^2 + r_{\eta,b}^2 + 2 r_{\eta,a} r_{\eta,b} r_{a,b}}{(1 - r_{a,b}^2) + \varepsilon}} \ , \qquad (14)$$

while ε is usually taken as a small value to prevent dividing by zero. The last λ_t (i.e., $t = 100ms$ at 100Hz sensor sampling rate system a queue that considers the last 10 observations) sets of the measured data have been considered for the joint correlation calculation. In other words, a queue with a size of λ_t is used for calculating the average in the correlation calculation. Based on the basic concept of the Pearson correlation coefficient [56], r_{χ_1, χ_2} between two sensor measurements (i.e., the measurement η and a in Eq(14)) can be defined as

$$r_{\chi_1, \chi_2} = \frac{\sum_{i=1}^{\lambda_t} (\chi_{1_i} - \overline{\chi}_1)(\chi_{2_i} - \overline{\chi}_2)}{\sqrt{\sum_{i=1}^{\lambda_t} (\chi_{1_i} - \overline{\chi}_1)^2} \sqrt{\sum_{i=1}^{\lambda_t} (\chi_{2_i} - \overline{\chi}_2)^2}} \ , \qquad (15)$$
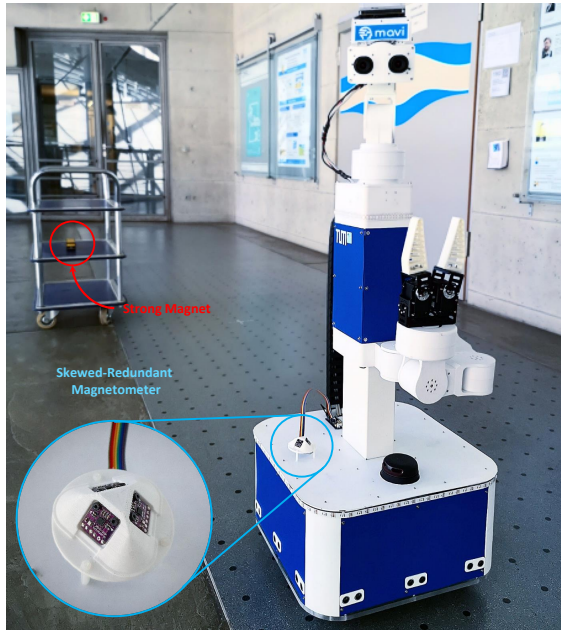
371

Fig. 7: Test setup of the skewed-redundant magnetometer platform in an indoor environment using the MAVI robot platform [57].

TABLE II: Root Mean Square Error (RMSE) of the proposed fusion system ($\lambda_t = 10$ and $\lambda_t = 50$) compared with the Kalman filtered aproach as well as the raw heading estimation. The measurement is a result of observation for over 100 iterations with and without external magnetic field perturbation.

Heading Estimation	No Perturbation	With External Perturbation	
Raw Estimation	16.79	63.41	(RMSE) [°]
Kalman Filter	12.28	36.65	
Proposed [$\lambda_t = 10$]	11.92	18.60	
Proposed [$\lambda_t = 50$]	**10.37**	**12.74**	

while the Pearson correlation coefficient is a measure of the linear correlation between two variables χ_1 and χ_2 and has a value between -1 and +1, where 1 is total positive linear correlation, 0 is no linear correlation, and -1 is total negative linear correlation. Further, the average values $\overline{\chi}_1$ and $\overline{\chi}_2$ are calculated from the last λ_t sets of the measured data. Lastly, considering Eq. 7, and taking the $\mathbf{B}_p^F$ as a magnetic field vector, the heading estimation can be extracted from Eq. 2.

V. Experimental Results

We designed a skewed-redundant magnetometer sensor platform and used the described correlation-based fusion to estimate the true Earth's magnetic field in an indoor environment. To evaluate the proposed method and mitigate the external magnetic field perturbation, a tetrahedron configuration is utilized and the azimuth angle estimation has been performed using the proposed fusion approach. We designed a 3D printed frame for mounting the sensors in a skewed-redundant structure. The designed sensing platform consists of three 3-axis AK8963C magnetometer sensors from Asahi Kasei Microdevices Corporation. The sensors are located on the planar surface of the 3D printed tetrahedron frame as shown in Fig. 6.

In order to evaluate the heading estimation and data recording, we used two different setups: one is a stationary platform mounted on an absolute rotary shaft encoder which can be rotated while providing the absolute orientation, and the second is the mobile robot platform MAVI [57] with the ability of self-localization using Inertial-LiDAR fusion (Fig. 7). In the first scenario, the computed heading from the turning encoder on a static table is used as a reference to assess the performance of the proposed method. We consider the offset of the heading reference that is calculated from the absolute shaft encoder to the Earth's true magnetic heading using the Earth's magnetic model [46].

The proposed fusion method has been tested in multiple scenarios portraying different magnetic configurations artificially introduced by permanent magnets or ferromagnetic materials. The comparison with the heading direction estimated from the raw and Kalman filtered data is shown in Table II. We used all the nine axes of the skewed-redundant magnetometer for the Kalman filter, but only one axis has been considered for naive heading estimation in the experiments. The result of the heading estimation (Fig. 8) using the correlation-based fusion shows the improvement and robustness in the presence of the external magnetic field anomaly. We achieved a RMSE of 12.74° in an indoor environment by using only the skewed-redundant magnetometer data fusion with no help of the gyroscopes. We also tested the proposed system in a large outdoor area with no additional permanent magnet or ferromagnetic material presence in the near surroundings. The result of this experiment shows that the output of the proposed fusion algorithm with no external disturbances provides a similar heading estimation accuracy as the Kalman filter approach. Furthermore, all the results are validated on an internally recorded dataset in our lab with more than 1.2 million sample sets which are encountered with and without external magnetic anomalies.

Additionally, we used the MAVI platform to perform a heading estimation using the Inertial-LiDAR localization system and used it as a reference for error calculation of the heading estimation in a diverse indoor environment. The experiment was conducted using a predefined trajectory followed by the MAVI platform while the estimated heading of the localization system is compared with the proposed fusion approach. In all the tests, the offset of the heading in the starting moment has been set to zero. As shown in Fig. 9, the average error encountered within the different environments is variable. This is caused by the changes in the external magnetic anomalies. The fluctuations observed for various indoor environments in the processed probability distributions indicate that the influences of magnetic disturbances on heading estimation are highly reliant on the construction of the building and
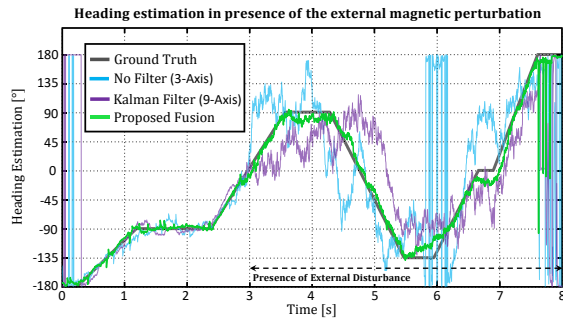
Fig. 8: Experimental result of the proposed method in comparison with no-filter magnetic sensor data and the Kalman filter for heading estimation in the presence of external magnetic field perturbations.
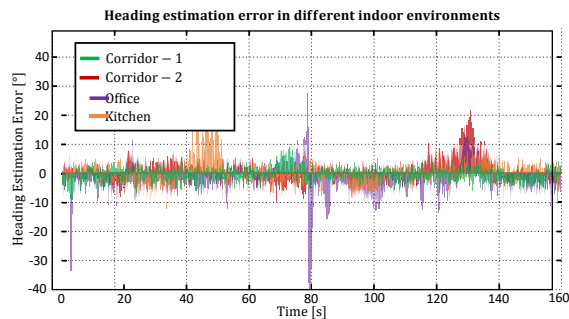

Fig. 9: Error of the heading estimation using the proposed method in different indoor environments in the presence of external magnetic field perturbations.

surroundings. Although the external anomalies were diverse from magnetic characteristics in tested indoor environments, the average error for azimuth estimation within these tests shows the reliability of the proposed approach.

VI. CONCLUSIONS

In this paper, we proposed a skewed-redundant magnetometer fusion approach for robust heading estimation in an indoor environment. We provided a description of heading estimation using the Earth's magnetic field and the different sources of perturbation for heading estimation in the presence of the external magnetic anomalies. We showed the principle of the Hall-effect magnetic sensors and analyzed the source of the external magnetic field on a 3-axis MEMS-based magnetic sensors. Inspired from the skewed-redundant configuration, we designed a tetrahedron platform using triple 3-axis magnetometer in order to detect and mitigate the magnetic perturbations. In order to filter the sensory data and mitigate the external disturbances, we proposed a correlation-based fusion approach for multi-magnetometer platforms. We further investigated the suggested method in different indoor environments and validated the algorithm using the internally

recorded dataset. Additionally, experiments using a mobile robot platform have been conducted in diverse scenarios to verify the generality of the proposed fusion method by means of robust heading estimation using the skewed-redundant magnetometer platform.

In future work, we plan to use a deep neural network to enhance the quality of the estimation. Additionally, in cases where there is a constructive disturbance along with changes in user dynamics, the perturbation mitigation approach will be further enhanced to take into account the uncertainty between perturbation and agent motion. Furthermore, the effect of the number of the sensing axes on the estimation of the true Earth's magnetic field needs to be studied using analytical models.

REFERENCES

[1] R. Harle, "A survey of indoor inertial positioning systems for pedestrians," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1281–1293, 2013.

[2] O. S. Salychev, *Applied Inertial Navigation: problems and solutions*. BMSTU press Moscow, Russia:, 2004.

[3] T. Aykut, M. Karimi, C. Burgmair, A. Finkenzeller, C. Bachhuber, and E. Steinbach, "Delay compensation for a telepresence system with 3d 360 degree vision based on deep head motion prediction and dynamic fov adaptation," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4343–4350, 2018.

[4] T. Aykut, C. Burgmair, M. Karimi, J. Xu, and E. Steinbach, "Delay compensation for actuated stereoscopic 360 degree telepresence systems with probabilistic head motion prediction," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2018, pp. 2010–2018.

[5] P. H. Mathisen and T. I. Fossen, "Robust navigation system for uavs in gnss-and magnetometer-denied environments," in *2019 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2019, pp. 1416–1424.

[6] T. Bailey, J. Nieto, J. Guivant, M. Stevens, and E. Nebot, "Consistency of the ekf-slam algorithm," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2006, pp. 3562–3568.

[7] J. Kelly and G. S. Sukhatme, "Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration," *The International Journal of Robotics Research*, vol. 30, no. 1, pp. 56–79, 2011.

[8] M. Karimi, S. S. Ghidary, R. Shekhar, T. D. Kane, and R. Monfaredi, "Magnetically anchored pan-tilt stereoscopic robot with optical-inertial stabilization for minimally invasive surgery," in *Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 10951. International Society for Optics and Photonics, 2019, p. 109511E.

[9] J. Zhang and S. 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.

[10] R. Mur-Artal and J. D. Tardós, "Visual-inertial monocular slam with map reuse," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 796–803, 2017.

[11] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale, "Keyframe-based visual–inertial odometry using nonlinear optimization," *The International Journal of Robotics Research*, vol. 34, no. 3, pp. 314–334, 2015.

[12] K. Gade, "The seven ways to find heading," *The Journal of Navigation*, vol. 69, no. 5, pp. 955–970, 2016.

[13] J. Wu, Z. Zhou, J. Chen, H. Fourati, and R. Li, "Fast complementary filter for attitude estimation using low-cost marg sensors," *IEEE Sensors Journal*, vol. 16, no. 18, pp. 6997–7007, 2016.

[14] M. Kok and T. B. Schön, "Magnetometer calibration using inertial sensors," *IEEE Sensors Journal*, vol. 16, no. 14, pp. 5679–5689, 2016.

[15] J. F. Vasconcelos, G. Elkaim, C. Silvestre, P. Oliveira, and B. Cardeira, "Geometric approach to strapdown magnetometer calibration in sensor frame," *IEEE Transactions on Aerospace and Electronic systems*, vol. 47, no. 2, pp. 1293–1306, 2011.

[16] M. Kok, J. D. Hol, T. B. Schön, F. Gustafsson, and H. Luinge, "Calibration of a magnetometer in combination with inertial sensors," in *2012 15th International Conference on Information Fusion*. IEEE, 2012, pp. 787–793.

[17] J. Lenz and S. Edelstein, "Magnetic sensors and their applications," *IEEE Sensors journal*, vol. 6, no. 3, pp. 631–649, 2006.

[18] M. Kok and T. B. Schön, "Maximum likelihood calibration of a magnetometer using inertial sensors," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 92–97, 2014.

[19] K. Han, H. Han, Z. Wang, and F. Xu, "Extended kalman filter-based gyroscope-aided magnetometer calibration for consumer electronic devices," *IEEE Sensors Journal*, vol. 17, no. 1, pp. 63–71, 2017.

[20] X. Tu, "Magnetometer calibration," May 6 2014, uS Patent 8,717,009.

[21] P. L. Worcester, D. J. Maxwell, and M. L. Trethewey, "Calibration of magnetic compasses," May 26 1992, uS Patent 5,117,375.

[22] J. Fang, H. Sun, J. Cao, X. Zhang, and Y. Tao, "A novel calibration method of magnetic compass based on ellipsoid fitting," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 6, pp. 2053–2061, 2011.

[23] I. Skog, "Inertial and magnetic-field sensor arrays-capabilities and challenges," in *2018 IEEE SENSORS*. IEEE, 2018, pp. 1–4.

[24] M. H. Afzal, V. Renaudin, and G. Lachapelle, "Assessment of indoor magnetic field anomalies using multiple magnetometers," in *ION GNSS*, vol. 10, 2010, pp. 21–24.

[25] V. Renaudin, M. H. Afzal, and G. Lachapelle, "Complete triaxis magnetometer calibration in the magnetic domain," *Journal of sensors*, vol. 2010, 2010.

[26] M. H. Afzal, V. Renaudin, and G. Lachapelle, "Multi-magnetometer based perturbation mitigation for indoor orientation estimation," *Navigation*, vol. 58, no. 4, pp. 279–292, 2011.

[27] ——, "Use of earth's magnetic field for mitigating gyroscope errors regardless of magnetic perturbation," *Sensors*, vol. 11, no. 12, pp. 11 390–11 414, 2011.

[28] V. Renaudin, M. H. Afzal, and G. Lachapelle, "New method for magnetometers based orientation estimation," in *IEEE/ION Position, Location and Navigation Symposium*. IEEE, 2010, pp. 348–356.

[29] G. Yin and L. Zhang, "Magnetic heading compensation method based on magnetic interferential signal inversion," *Sensors and Actuators A: Physical*, vol. 275, pp. 1–10, 2018.

[30] S. Guerrier, "Improving accuracy with multiple sensors: Study of redundant mems-imu/gps configurations," in *Proceedings of the 22nd international technical meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2009)*, 2009, pp. 3114–3121.

[31] X. Zhang, L. Xiao, C. Zou, L. Guo, Y. Jin, and G. Shi, "A novel calibration method of electronic compass based on multi-redundancy," in *2017 IEEE International Conference on Real-time Computing and Robotics (RCAR)*. IEEE, 2017, pp. 156–161.

[32] I. Colomina, M. Giménez, J. Rosales, M. Wis, A. Gómez, and P. Miguelsanz, "Redundant imus for precise trajectory determination," in *Proceedings of the 20th ISPRS Congress, Istanbul, Turkey*, vol. 1223. Citeseer, 2004, p. 17.

[33] S. Sukkarieh, P. Gibbens, B. Grocholsky, K. Willis, and H. F. Durrant-Whyte, "A low-cost, redundant inertial measurement unit for unmanned air vehicles," *The International Journal of Robotics Research*, vol. 19, no. 11, pp. 1089–1103, 2000.

[34] R. Giroux, S. Sukkarieh, and M. Bryson, "Implementation of a skewed-redundant low-cost ins in a fast-prototyping environment," in *Proceedings of the Institute of Navigation National Technical Meeting*, 2004.

[35] M. Jafari, "Optimal redundant sensor configuration for accuracy increasing in space inertial navigation system," *Aerospace Science and Technology*, vol. 47, pp. 467–472, 2015.

[36] A. Osman, B. Wright, A. Noureldin, and N. El-Sheimy, "Multi-sensor inertial navigation systems employing skewed redundant inertial sensors," in *ION GNSS 19th International Technical Meeting of the Satellite Division*, 2006, pp. 26–29.

[37] Y. Yu and M. Cheng, "Comparative research of redundant strap down inertial navigation system based on different configuration schemes," in *AIP Conference Proceedings*, vol. 1967, no. 1. AIP Publishing, 2018, p. 040026.

[38] H. Naseri and M. Homaeinezhad, "Improving measurement quality of a mems-based gyro-free inertial navigation system," *Sensors and Actuators A: Physical*, vol. 207, pp. 10–19, 2014.

[39] S. Ghasemi-Moghadam and M. Homaeinezhad, "Attitude determination by combining arrays of mems accelerometers, gyros, and magnetometers via quaternion-based complementary filter," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 31, no. 3, p. e2282, 2018.

[40] M. W. Givens and C. Coopmans, "A survey of inertial sensor fusion: Applications in suas navigation and data collection," in *2019 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2019, pp. 1054–1060.

[41] F. Liu, Z. Su, H. Zhao, Q. Li, and C. Li, "Attitude measurement for high-spinning projectile with a hollow mems imu consisting of multiple accelerometers and gyros," *Sensors*, vol. 19, no. 8, p. 1799, 2019.

[42] A. Grosz, M. J. Haji-Sheikh, and S. C. Mukhopadhyay, *High sensitivity magnetometers*. Springer, 2017.

[43] R. Hansenw, R. L. Kleinberg, and A. Kaufman, *Principles of the Magnetic Methods in Geophysics, Volume 42 (Methods in Geochemistry and Geophysics)*. Elsevier, 2009.

[44] S. Tumanski, *Handbook of magnetic measurements*. CRC press, 2016.

[45] E. Ramsden, *Hall-effect sensors: theory and application*. Elsevier, 2011.

[46] "Global geomagnetic models," 2019. [Online]. Available: https://geomag.bgs.ac.uk/research/modelling/modelling.html

[47] M. Sturza, "Skewed axis inertial sensor geometry for optimal performance," in *Digital Avionics Systems Conference*, 1988, p. 3874.

[48] S. Qiu, Z. Wang, H. Zhao, K. Qin, Z. Li, and H. Hu, "Inertial/magnetic sensors based pedestrian dead reckoning by means of multi-sensor fusion," *Information Fusion*, vol. 39, pp. 108–119, 2018.

[49] H. Liu, R. Shor, and S. S. Park, "Intelligent filter for accurate subsurface heading estimation using multiple integrated mems sensors," in *2018 IEEE SENSORS*. IEEE, 2018, pp. 1–4.

[50] H. Hou, *Modeling inertial sensors errors using Allan variance*. University of Calgary, Department of Geomatics Engineering, 2004.

[51] Z. Xing and D. Gebre-Egziabher, "Modeling and bounding low cost inertial sensor errors," in *2008 IEEE/ION Position, Location and Navigation Symposium*. IEEE, 2008, pp. 1122–1132.

[52] G. Troni and R. M. Eustice, "Magnetometer bias calibration based on relative angular position: Theory and experimental comparative evaluation," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2014, pp. 444–450.

[53] B. Dil, G. Hendeby, F. Gustafsson, and B. J. Hoenders, "Approximate diagonalized covariance matrix for signals with correlated noise," in *2016 19th International Conference on Information Fusion (FUSION)*. IEEE, 2016, pp. 521–527.

[54] D. Unsal and K. Demirbas, "Estimation of deterministic and stochastic imu error parameters," in *Proceedings of the 2012 IEEE/ION Position, Location and Navigation Symposium*. IEEE, 2012, pp. 862–868.

[55] Y. Stebler, S. Guerrier, and J. Skaloud, "An approach for observing and modeling errors in mems-based inertial sensors under vehicle dynamic," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 11, pp. 2926–2936, 2015.

[56] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in *Noise reduction in speech processing*. Springer, 2009, pp. 1–4.

[57] M. Karimi, T. Aykut, and E. Steinbach, "Mavi: A research platform for telepresence and teleoperation," *arXiv preprint arXiv:1805.09447*, 2018.

Chapter 4

Deep Fusion of a Skewed Redundant Magnetic and Inertial Sensor for Heading State Estimation in a Saturated Indoor Environment

4.1 Attitude and Heading Reference Estimation based on Skewed Redundant MIMU Deep Fusion

It is indispensable for various robot applications to be able to detect, estimate and make use of the attitude and heading in an indoor environment with respect to a known reference. Affordable Attitude and Heading Reference System (AHRS) use low-cost solid-state MEMS-based sensors in order to provide accurate attitude and heading information. With such a system, the accuracy of heading estimation is typically degraded because of the drift that occurs from the gyro measurements as well as the distortion induced by the Earth's magnetic field. In this chapter, we present a novel approach for robust indoor heading and attitude estimation using skewed redundant inertial and magnetic sensors. The presented approach uses Recurrent Neural Networks (RNN) fusion as a suitable technique for performing robust heading estimation as well as compensating for anomalies in the external magnetic field. We employ our previously described correlation-based filter model both for preprocessing the data and for enabling perturbation mitigation. In experiments we have conducted using the proposed method, we demonstrate that the proposed method enables the mitigation of anomalies in a saturated indoor environment and achieve accurate and robust attitude and heading estimation over the long run. Result of proposed technique as a robust attitude and heading estimation is used in the next chapter where we tackle the SLAM problem.

4.2   Authors Contribution

Design and development of the measurement method using tetrahedron-shaped platform, a deep-fusion-based technique, and execution of the experiments were performed by Mojtaba Karimi. Data visualization were conducted by Edwin Babaians. Recording and labeling the data as well as the printed circuit board was conceptualized by Mojtaba Karimi. Martin Oelsch, and Mojtaba Karimi evaluated the experimental results. Eckehard Steinbach supervised this work. The original draft was written by Mojtaba Karimi. The paper was reviewed and edited by all authors.

Title:

Deep Fusion of a Skewed Redundant Magnetic and Inertial Sensor for Heading State Estimation in a Saturated Indoor Environment

Authors:
Mojtaba Karimi, Edwin Babaians, Martin Oelsch, and Eckehard Steinbach

World Scientific
www.worldscientific.com

Deep Fusion of a Skewed Redundant Magnetic and Inertial Sensor for Heading State Estimation in a Saturated Indoor Environment

Mojtaba Karimi*, Edwin Babaians†, Martin Oelsch‡
and Eckehard Steinbach§

Department of Electrical and Computer Engineering
Technical University of Munich, 80333 Munich, Germany
*mojtaba.karimi@tum.de
†edwin.babaians@tum.de
‡martin.oelsch@tum.de
§eckehard.steinbach@tum.de

Robust attitude and heading estimation in an indoor environment with respect to a known reference are essential components for various robotic applications. Affordable Attitude and Heading Reference Systems (AHRS) are typically using low-cost solid-state MEMS-based sensors. The precision of heading estimation on such a system is typically degraded due to the encountered drift from the gyro measurements and distortions of the Earth's magnetic field sensing. This paper presents a novel approach for robust indoor heading estimation based on skewed redundant inertial and magnetic sensors. Recurrent Neural Network-based (RNN) fusion is used to perform robust heading estimation with the ability to compensate for the external magnetic field anomalies. We use our previously described correlation-based filter model for preprocessing the data and for empowering perturbation mitigation. Our experimental results show that the proposed scheme is able to successfully mitigate the anomalies in the saturated indoor environment and achieve a Root-Mean-Square Error of less than 2.5° for long-term use.

Keywords: Heading estimation; skewed redundant; RNN; magnetic and inertial sensor.

1. Introduction

Attitude and heading estimation are fundamental requirements for robotics, human–machine interaction, and navigation in indoor environments [1–3]. The ability of self-localization for autonomous systems (e.g. unmanned aerial vehicles, mobile robots) or observing head motion for teleportation or virtual reality applications [4, 5], crucially depend on the reliable estimation of attitude and heading angles [6–8]. The result of camera-based or LiDAR-based SLAM systems, for instance, becomes more reliable by utilizing robust attitude and heading data, which can be attained mainly from an Attitude and Heading Reference System (AHRS). Similarly, a Magnetic and Inertial Measurement Unit (MIMU) can be used for tightly coupled fusion approaches [9–11]. The main difference between AHRSs in contrast to MIMUs is the preprocessing of the sensory data in order to provide absolute orientation estimation

314   *M. Karimi et al.*

with respect to the Earth's geometrical characteristics. The known reference of AHRS comes from the Earth's gravitational and magnetic field. In this regard, attitude estimation refers to the angular measurement on the vertical plane, concerning the local level frame, and computed as inclinations. Similarly, angular incline on the horizontal plane, with respect to the actual geographical north, is referred to as heading angle (also known as yaw or azimuth) [12].

Affordable AHRSs are typically using small-scale sensors based on solid-state Micro-Electro-Mechanical Systems (MEMS). The MEMS-based sensors have the advantages of low cost, low power consumption, and high update rate. A general 9-DoF AHRS consists of a 3-DoF accelerometer, a 3-DoF gyroscope, and a 3-DoF magnetic field meter. Exploiting the complementary nature of the MEMS-based inertial sensors, 3-DoF accelerometers along with 3-DoF gyroscopes are mainly fused for attitude estimation [13]. Similarly, the MEMS-based AHRSs mostly use a single 3-DoF magnetometer to measure the magnetic field strength and utilize the previously determined relative calibration values to extract the Earth's geographic azimuth. In practice, the magnetometer-based azimuth estimation is fused with the gyroscopes to provide a more stable heading estimation. Nonetheless, the general accuracy for heading estimation on such an AHRS highly depends on the Earth's true magnetic field measurement in the long term heading estimation. This is mainly because of the accumulated drift caused by the integration of the gyro data, and this drift is conventionally compensated using the heading estimation with magnetometers.

Determining error-free and reliable heading angle with respect to a known reference is problematic in the case of indoor applications. This problem is mainly because of different sources of errors in the MEMS-based magnetometer measurements [14–17]. The most dominant source of error for such a magnetometer sensor is the presence of magnetic sources in an indoor environment. Permanent magnets and ferromagnetic materials are the preeminent cause of such magnetic perturbation. In this regard, calibration of the MEMS-based magnetometers for heading estimation has been investigated thoroughly during the last decades [14, 18]. Probabilistic models for sensor calibration and fusion such as the Extended Kalman Filter (EKF), the Unscented Kalman Filter (UKF), and Ellipsoid Fitting are well known in this field [19–22]. However, dynamic magnetic deviation caused by external anomalies cannot be fully compensated using the sensor calibration methods [23]. To overcome the aforementioned problem, redundant sensor systems have shown promising results as they are using redundant observations to determine the source of error and increase the estimation accuracy [24–26].

Multiple magnetometers combined with gyroscopes were used to approximate the external magnetic disturbances and provide perturbation-free heading estimation [25, 27, 28]. In these works, the authors used an estimation of the expected magnetic field using the temporary rotation vector to compensate for the external magnetic perturbation. These approaches need a precise and simultaneous calibration of the magnetometers and the gyroscopes to be able to provide an

acceptable heading estimation. In [29], a magnetic heading compensation method for disturbances introduced by near-surface interference was proposed. The authors used inverse magnetic anomaly vector estimation to compensate for the external perturbation. The main limitations of this approach are the predefined magnetic field model and one-dimensional perturbation mitigation. In general, these approaches are usually bulky, expensive, and often rely on expensive calibration procedures.

Nonetheless, it has been shown that the accuracy of the magnetic field measurement is significantly improved by using redundant magnetometer sensors [30–33]. However, the non-deterministic errors cannot be completely removed in real application scenarios due to their transient nature. This is mainly because of limitations in the system identification procedure, nonlinear dynamic errors, and the use of strict calibration models. The skewed redundant inertial sensor fusion, on the other hand, provides a more reliable output, as shown in [34–37]. The skewed configuration for redundant sensor geometry has been widely investigated for attitude estimation by applying this method on the observations from inertial-only sensors (e.g. accelerometers and gyroscopes). However, this method has been poorly investigated for redundant magnetometers.

Inspired by the success of redundant integrated sensors for processing high-dimensional sensor data [25, 38, 39], and the benefits of redundant, non-orthogonal sensor configurations [39–42], we propose a skewed redundant magnetometer and inertial platform for perturbation-free heading estimation. Following our prior work [43], the approach undertaken in this paper relies on the use of redundant sensor observations to enhance the overall accuracy of the heading estimation while ensuring that the latency and the update rate of the estimation remains acceptable for teleoperation and autonomous robotic applications. Unique non-orthogonal placement of triple 3-axis MIMUs is designed based on the Hall-effect sensor principle to mitigate the external magnetic perturbation by extracting the saturated signals (Fig. 1). The proposed approach uses a correlation-based filter model for magnetic field observation and recurrent neural network (RNN)-based sensor fusion for external magnetic perturbation mitigation and robust heading estimation. The skewed redundant sensor platform and deep fusion model show substantial improvement in heading estimation accuracy in an indoor environment. Our contributions can be summarized as follows:

- We introduce a skewed redundant magnetic and inertial sensor platform and the methodology to configure redundant sensing kinematic for non-orthogonal observation.
- We propose a correlation-based preprocessing filter and investigate multiple RNN-based deep fusion models to improve accuracy and mitigate the external perturbation for robust heading estimation in a saturated indoor environment.
- Multiple experiments were designed and conducted to evaluate the proposed system's performance toward the reliable state estimation in different indoor
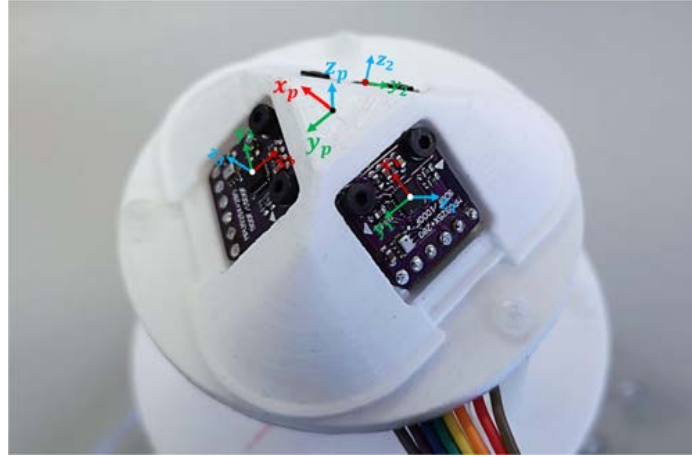
316   *M. Karimi et al.*

Fig. 1.   Skewed redundant magnetic and inertial sensor for robust heading estimation in indoor environments using RNN-based fusion.

environments with the presence of external magnetic anomalies. We compared the result of the heading estimation to multiple well-known and state-of-the-art fusion methods.

In this work, we extend our previously presented correlation-based skewed redundant magnetic fusion concept [43] by incorporating an inertial-based multi-sensor fusion for robust heading estimation using RNNs. Considering system dynamics and saturated observations, our proposed approach decrease the external perturbations' effect using the hybrid filter model and the deep fusion. In doing so, we are able to momentarily extract the error patterns and fuse MIMUs, which results in a more robust heading estimation for long-term use.

The rest of this paper is organized as follows. Section 2 describes the magnetic field components and the basics of heading estimation using magnetic hall-sensors. Section 3 introduces the skewed redundant sensor platform and the kinematic of the non-orthogonal observation, while Sec. 4 details the correlation-based preprocessing filter and RNN-based fusion models. Experimental evaluations are presented in Sec. 5. Finally, Sec. 6 concludes this paper and examines the limitations and potential future works.

2. Magnetic Field Components and Heading Estimation

A combination of a 3-axis accelerometer, a 3-axis gyroscope, and a 3-axis magnetometer is typically used in low-cost AHRSs. The estimated angles are defined as the corresponding Euler angles roll (α), pitch (β), and yaw (γ). In order to improve the robustness of the roll and pitch estimation, the measurements provided by the gyroscope and accelerometer are typically fused. The magnetometer, on the other hand, can be complemented using the gyroscope for heading angle estimation. Heading

angle estimation with respect to a known reference can be performed by considering
the fact that solid-state magnetometers are used to measure changes in the strength
of the Earth's magnetic field. Dominant technologies for the solid-state MEMS-based
magnetometers are based on the Hall-effect, or the Magneto Resisting effect [17, 44].
Considering that the magnetic observation is a vector field, it is characterized by its
magnitude and direction or its components along the coordinate axes. A sensor with
at least three axes is required to determine the Earth's magnetic field vector (see
Fig. 2). In a Cartesian coordinate system at the measurement point p, the magnetic
field vector $\mathbf{b}_p \in \mathbb{R}^3$ can be defined as

$$\mathbf{b}_p = b_{p,x}\boldsymbol{i} + b_{p,y}\boldsymbol{j} + b_{p,z}\boldsymbol{k}, \tag{1}$$
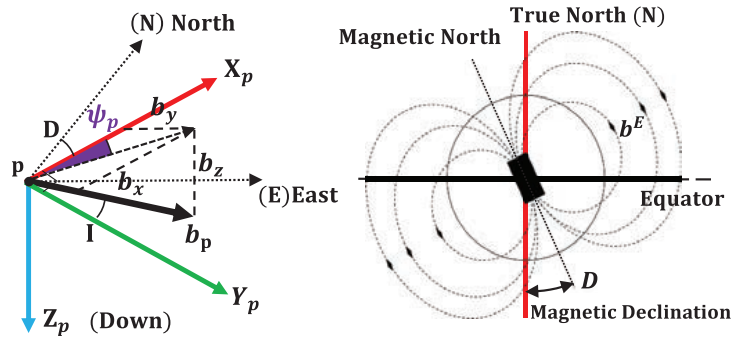
Fig. 2.   Components of the Earth's magnetic field and the abstraction of the heading estimation using a
3-axis magnetometer. The angle ψ_p is the true heading with respect to horizontal plane, with respect to the
actual geographical north.

where the x-axis is oriented along the geographical meridian, and the direction to the
north is positive, the y-axis along the parallel with a positive direction toward the east,
and the z-axis is directed downward. The observation point p is the origin of the
measurement coordinate system. The angle $\mathbf{I}$ between the horizontal plane and the
vector $\mathbf{b}_p$ is called the inclination [45]. The local magnetic declination $\mathbf{D}$, sometimes
called magnetic variation, is the angle between the magnetic north and the true north.
Accordingly, the Earth's geographic azimuth ψ_p can be derived by considering the local
declination $\mathbf{D}$ and the perpendicular components of the measurement vector $\mathbf{b}_p$ as

$$\psi_p = \tan^{-1}\left(\frac{\mathbf{b}_{p,y}}{\mathbf{b}_{p,x}}\right)\mathbf{D}. \tag{2}$$

The magnetic field of a dipole, similarly, is characterized by its magnetic moment
$\mathbf{M}$. The strength and direction of this field depend not only on the magnitude of the
magnetic moment but also on its orientation relative to the direction of the magnetic
field [45]. In the presence of multiple dipoles, the magnetic field $\mathbf{M}_g$ is defined by the
sum of all magnetic moments $\mathbf{M}_i$ and is given by

$$\mathbf{M}_g = \sum \mathbf{M}_i. \tag{3}$$

318   *M. Karimi et al.*

The curl of the magnetic vector potential, the magnetic field $\mathbf{b}_p$, is governed by the density of magnetic moments at the measurement point and is derived classically using the Biot–Savart law [47]. The combined field can change the measurement field vector $\mathbf{b}_p$ due to the presence of multiple dipoles. This change in the magnetic field is the leading cause of heading errors in an indoor environment. As shown in Fig. 3, the
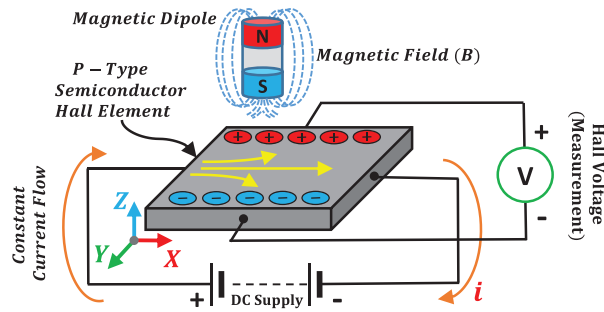
Fig. 3.   Single-axis Hall-effect sensor principle [46]. The output signal from a Hall-effect sensor is a function of the magnetic field density around the device. A minimum of three perpendicular axes is needed for the magnetic field vector measurement.

effect of the external magnetic field on the p-type Hall element depends highly on the orientation and the distance of the external magnetic moment to the perpendicular sensing plane [46]. Moreover, a minimum number of three Hall elements perpendicular to each other is required to determine the magnetic field $\mathbf{b}_p$ as it is shown in Eq. (1). This implies that the measuring axis with the lowest angle to the external magnetic anomaly will report a non-correlated measurement (generally saturated observation) compared to the other axes. Additionally, in a perturbation-free environment, the norm of the magnetic field vector measurement should be equal to the magnitude of the Earth's magnetic field. This norm, H_p can be extracted from the Earth's geomagnetic model [48]. Besides, these observations have a complex pattern in their essence and lead to a nonlinear relation to a true heading estimation. The aforementioned principles are the main motivation of proposing the skewed redundant sensing platform and RNN-based fusion.

3. Skewed Redundant Platform

Data measured by sensors are inherently clouded by noise; that is to say, they display a band of measurements around the actual signal. Methods have been developed to reduce the uncertainty around the measurement such that there is higher confidence in the measured signal. A significant outcome in [49] shows that an ideal setup of redundant inertial sensors in terms of fault tolerance is a non-orthogonal configuration in which the delicate axes of each triaxial sensor are not aligned, and transformations are performed to align them in a virtual coordinate frame. This method has been examined on inertial sensors such as accelerometers as well as gyroscopes and showed a significant improvement in the measurement accuracy for attitude and

heading estimation [34, 36, 37]. Following the non-orthogonal inertial sensor configuration concept, we propose redundant sensing kinematic for non-orthogonal observation based on skewed redundant placement.

As we described in Sec. 2, considering the principle of the hall-effect magnetometers, it is observable that the perpendicular axis to the external magnetic anomalies can be affected dramatically [43]. The pattern of this change in the data is recognizable, and this motivates us to examine different mounting structures with redundant sensor setup. It is difficult to determine which configuration yields the best performance since many possible arrangements make it challenging to derive a deterministic relation between the system errors and the configuration parameters. However, we investigated several configurations and analyzed the data to determine for which of them the external magnetic disturbance is perceivable. Additionally, we considered the minimum number of sensors and the coverage orientation. The skewed configuration in comparison to the orthogonal structures showed significantly determinable results. Using a tetrahedron skewed configuration [49], we recorded the data shown in Fig. 4, which illustrates that the external magnetic perturbation causes a determinable pattern in the observation using non-orthogonal configuration. As the plot shows, the axis, which has the minimum angle difference to the external magnetic anomalies source, reports a saturated output. In contrast, the
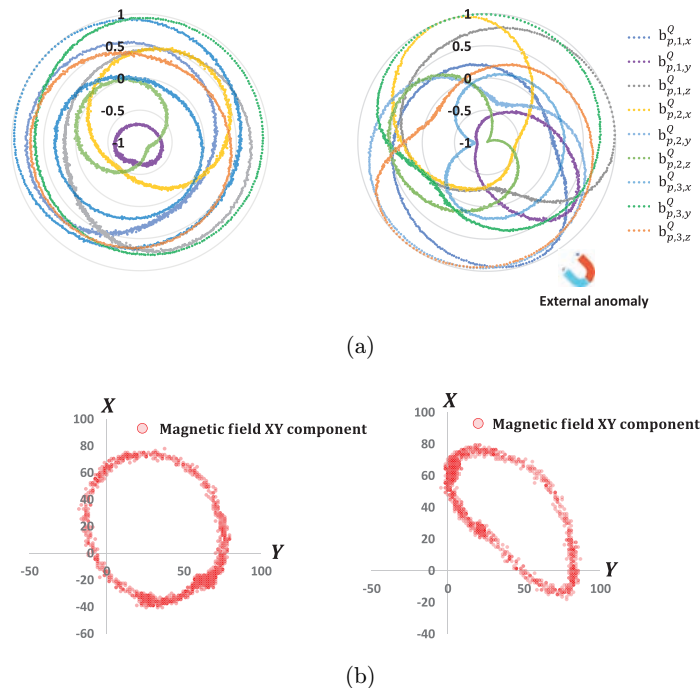
(a)

(b)

Fig. 4.   Output signals from a skewed redundant magnetometer sensor at measurement point p and the effect of the external disturbance on the specific sensing axes, which are perpendicular or have a small angle to the source of the external magnetic moment. (a) Shows the normalized magnetic field observations. (b) The extracted X and Y components of the magnetic field.

320   *M. Karimi et al.*

other axes of the magnetic measurements are less impacted. This causes the
extracted magnetic field components to face a dynamic soft-iron error.

The geometrical relationship of the skewed redundant MIMU configuration was
leveraged to develop a method that extended the dynamic range of the sensors. Three
sensor frames were configured in the designed sensing platform, each having its own
orthogonal three-dimensional axes (see Fig. 5). The relationship between the skewed
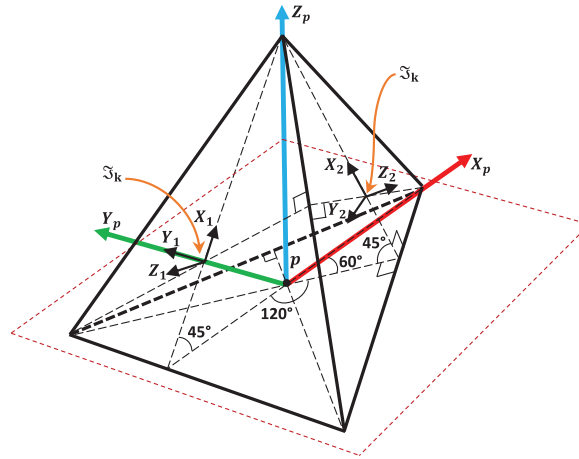
Fig. 5.   Tetrahedron configuration model of the skewed redundant magnetometer platform using three 3-
axis magnetometer sensors. The virtual coordinate frame for the measurement is shown with the center
point p, and only two sensor frames are depicted for simplification.

sensors was defined through coordinate transformations. To navigate between
skewed reference frames, rotational matrices were used to go between each coordi-
nate frame. Coordinate transformations done for practical implementation are more
easily visualized with Euler angles. Euler angles refer to a yawing orientation, a
pitching orientation, and a rolling orientation. Defining the principal axis as yaw for
the z-axis, pitch for the y-axis, and roll for the x-axis relates the Euler angles to the
axis of motion as shown in Fig. 6. Euler angles enable an intuitive understanding of
the attitude and heading changes within the body frame.

3.1. *Magnetic field measurement*

To determine the Earth's true magnetic field, we consider the measured magnetic
field as a combination of the external magnetic perturbation and the Earth's mag-
netic field. Taking Eq. (3) into consideration, the measured magnetic field $\mathbf{b}^Q \in \mathbb{R}^3$
in the sensor frame with center point $\Im$ can be defined as

$$\mathbf{b}_\Im^Q = \mathbf{b}_\Im^E + \mathbf{b}_\Im^\epsilon, \tag{4}$$

where $\mathbf{b}_\Im^\epsilon$ is the external magnetic field disturbance and $\mathbf{b}_\Im^E$ is the Earth local
magnetic field. To have all the true signals in the measurement point p, sensor data
needs to be aligned to a defined virtual coordinate within the body frame.
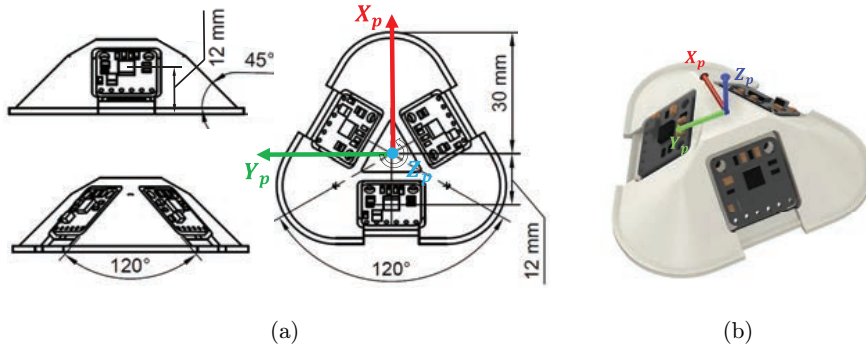
54

Chapter 4. Deep Fusion of a Skewed Redundant Magnetic and Inertial Sensor for Heading State Estimation in a Saturated Indoor Environment

(a) (b)

Fig. 6. Structure of the 3D printed model of the tetrahedron platform with the illustration of the dimension of the module.

Additionally, we normalize the measured magnetic field vector to be able to use the correlation-based features later for sensor filtering and fusion. We can define $\mathbf{b}_{p,k}^{Q}$ for each sensor k using

$$\mathbf{b}_{p,k}^{Q} = \mathbf{R}_{\Im,k}^{p} \left[\frac{b_{\Im,k,x}^{Q}}{\|\mathbf{b}_{\Im,k}^{Q}\|} \quad \frac{b_{\Im,k,y}^{Q}}{\|\mathbf{b}_{\Im,k}^{Q}\|} \quad \frac{b_{\Im,k,z}^{Q}}{\|\mathbf{b}_{\Im,k}^{Q}\|} \right]^{T}, \tag{5}$$

where $\mathbf{R}_{\Im,k}^{p} \in SO(3)$ is the Direction Cosine Matrix (DCM) with respect to roll (α_k), pitch (β_k), and yaw (γ_k) angles. The angles roll, pitch, and yaw are the representative rotation of the 3-axis magnetometer from the sensor coordinate frame with the center point $\Im$ around the virtual coordinate frame with the center point p. Based on the skewed configuration, the proposed structure with the mounting angles is depicted in Fig. 5. A Direction Cosine Matrix (DCM) can be defined for an individual sensor k according to the static angles using

$$\mathbf{R}_{\Im,k}^{p} = \begin{pmatrix} c\beta_k c\gamma_k & -c\beta_k s\gamma_k & s\gamma_k \\ c\alpha_k s\gamma_k + s\alpha_k s\beta_k s\gamma_k & c\alpha_k c\gamma_k - s\alpha_k s\beta_k s\gamma_k & -s\alpha_k c\beta_k \\ s\alpha_k s\gamma_k - c\alpha_k s\beta_k c\gamma_k & s\alpha_k c\gamma_k + c\alpha_k s\beta_k s\gamma_k & c\alpha_k c\beta_k \end{pmatrix}. \tag{6}$$

3.2. *Inertial measurement*

Similar to the magnetic field vectors, accelerometer and gyroscope data are measured in the same coordinate frame $\Im$. This means we can use the same DCM as defined in Eq. (6). Accordingly, the observed acceleration $\mathbf{a}_{p,k} \in \mathbb{R}^3$ in the body frame with center point p can be defined by

$$\mathbf{a}_{p,k} = \mathbf{R}_{\Im,k}^{p} [\, a_{\Im,k,x} \quad a_{\Im,k,y} \quad a_{\Im,k,z} \,]^{T}, \tag{7}$$

where $\mathbf{a}_{\Im,k}$ is the acceleration vector for sensor k in m/s^2. Likewise, the angular rotation $bfv_{p,k} \in \mathbb{R}^3$ in the body frame with center point p is defined by

$$\mathbf{v}_{p,k} = \mathbf{R}_{\Im,k}^{p} [\, v_{\Im,k,x} \quad v_{\Im,k,y} \quad v_{\Im,k,z} \,]^{T}, \tag{8}$$

where $\mathbf{v}_{\Im,k}$ is the angular velocity vector in $\mathrm{rad} \cdot \mathrm{s}^{-1}$.

322 *M. Karimi et al.*

4. Skewed Redundant RNN Fusion

To the best of our knowledge, this is the first work that applies deep learning for heading estimation using a skewed redundant sensor platform. For that reason, we investigated numerous structures ranging from simple to more sophisticated dense models. However, the first step is to preprocess the sensory data and use the filtered observation to train the deep model (see Fig. 7). We use a correlation-based filter for
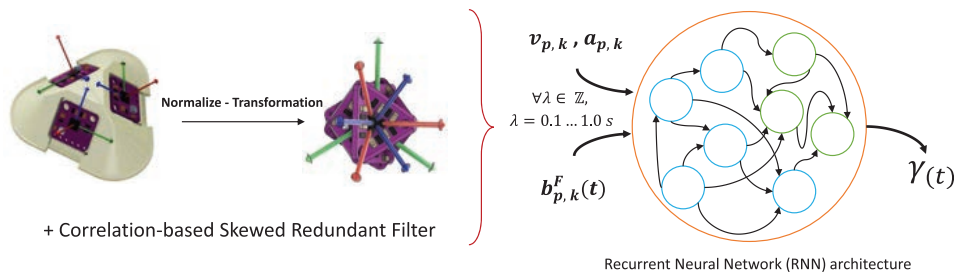
Fig. 7.   System overview of the tetrahedron platform with the illustration of the skewed redundant RNN-based fusion.

magnetic field observation that allows reducing the external magnetic perturbation. In addition, we feed the raw inertial data (e.g. acceleration and the angular velocity) to the network in order to compensate for the dynamics of the system and estimate a robust heading for long-term performance in an indoor environment.

4.1. *Correlation-based filter model*

The main sources of magnetometer errors are the sensor offsets, scale factor, non-orthogonality, and magnetic deviation [25, 50, 51]. The latter is based on the sensor's magnetization characteristics and can have a static or dynamic source. The static magnetic deviation error consists of soft- and hard-iron parts. In this paper, we are not going to the details of the sensor hard-iron and soft-iron calibration procedure. We assume the measured magnetic field $\mathbf{b}_{p,k}^{Q}$ is calibrated by using the method described in [14]. Additional to the static magnetic deviation error, which can be addressed by applying the offline calibration, the dynamic magnetic deviation error is still problematic and subsists as the most dominant error source. Therefore, the proposed multi-magnetometer filter model uses dynamic weighting based on the normalized correlation coefficient analysis to refine the axes affected by the external disturbance. This mitigates the external dynamic magnetic field perturbation and allows for the estimate of the true Earth's magnetic field. Let us define the true Earth's magnetic field vector $\mathbf{b}_{p}^{F} \in \mathbf{R}^{3}$ as

$$\mathbf{b}_{p,k}^{F} = \mathbf{J}_k \mathbf{M}_k (\mathbf{S}_{p,k} \mathbf{W}_{p,k}(t) \mathbf{b}_{p,k}^{Q} + \mathbf{c}_{p,k}), \tag{9}$$

where $n = 3$ is the number of sensed magnetic field vectors that are transformed to the measurement point p in the skewed-redundant multi-magnetometer setup, $\mathbf{S}_{p,k}$ is

the scale factor of the sensor k, $\mathbf{c}_{p,k} \in \mathbb{R}^3$ is the bias vector, $\mathbf{W}_p(t) \in \mathrm{SO}(3)$ is the white noise diagonal matrix, $\mathbf{M}_k \in \mathrm{SO}(3)$ represents the skew-symmetric misalignment matrix, and $\mathbf{J}_k \in \mathrm{SO}(3)$ is the adaptive correlation-based diagonal dynamic weighting matrix.

We use the general definition of the $\mathbf{S}_{p,k}$ and $\mathbf{M}_k$ for the proposed magnetometer sensor setup from [26, 31]. It should be noted that separating all components of the encountered errors for the magnetometer platform requires unique setups (e.g. turning tables) and is hard to accomplish. However, this is not essential in the case of the MEMS-based sensors because some error elements predominate. The error model, therefore, can be simplified [52, 53]. The biases $\mathbf{c}_{p,k}$ are approximated for each 3-axis magnetometer sensor and can be defined as

$$\mathbf{c}_{p,k} = \mathbf{c}_k^{\mathrm{C}} + \mathbf{c}_k^{\mathrm{RC}} + \mathbf{c}_k^{\mathrm{GM}}(t) + \mathbf{c}_k^{\mathrm{RW}}(t) + \mathbf{c}_k^{\mathrm{BI}}(t), \tag{10}$$

where $\mathbf{c}_k^{\mathrm{C}}$ is the constant component of the bias vector and can be extracted from sensor design characteristics, $\mathbf{c}_k^{\mathrm{RC}}$ is the stochastic process with maximum constant value which is extracted from the hard-iron calibration procedure and having the distribution $\mathbf{c}_k^{\mathrm{RC}} \sim N(\mu_k^{\mathrm{c}}, \sigma_k^{\mathrm{c}^2})$. The $\mathbf{c}_k^{\mathrm{GM}}$ is a first-order Gauss–Markov stochastic process applied on the sensor constant bias estimation. Constant biases of the sensors are extracted from the standard deviation of the measurements considering that each measurement is a discrete sample measurement at t [54]. Therefore, $\mathbf{c}_k^{\mathrm{GM}}$ is defined as

$$\mathbf{c}_k^{\mathrm{GM}}(t) = \zeta \cdot \dot{\mathbf{c}}_k^{\mathrm{GM}}(t) + \mathbf{E}_k(t), \tag{11}$$

where $\zeta = 1/T$ is the inverse of the correlation time with the initial condition

$$\lim_{t_0 \to \infty} \mathbf{c}_k^{\mathrm{GM}}(t_0) = 0,$$

and $\mathbf{E}_k(t)$ is a zero-mean Gaussian white-noise process which is applied on the sensor constant bias vector. We denote the finite bandwidth of the respective magnetometer data by Δf, which can be calculated from the auto-correlation function as expressed in [55]. The Nyquist–Shannon sampling rate equals $2\Delta f$ which is equal to two samples per period T. Hence, defining $T/2$ as the correlation time, the Nyquist sampling rate ζ is equal to the inverse of the correlation time. The $\mathbf{c}_k^{\mathrm{RW}}(t)$ is a random walk stochastic process defined as

$$\dot{\mathbf{c}}_k^{\mathrm{RW}}(t) = \mathbf{E}_k(t)$$

with the initial condition $\mathbf{c}_k^{\mathrm{RW}}(t_0) = 0$. The $\mathbf{c}_k^{\mathrm{BI}}(t)$ are the bias instabilities which were added to the original model developed in [30, 56] to describe more precisely the noise characteristics observed for the skewed-redundant magnetometer platform (see Fig. 5). We use the same definition of the proposed bias instabilities vector as in [30]. The white noise process $\mathbf{W}_{p,k}$ is defined as

$$\mathbf{W}_{p,k}(t) = \sigma_{p,k}^{\mathrm{w}}\mathbf{E}_k(t)\boldsymbol{I}_3, \tag{12}$$

324   *M. Karimi et al.*

where $\sigma_{w,k}^{\mathrm{w}}$ is the standard deviation of the measurements, and $\boldsymbol{I}_3 \in \mathbb{R}^3$ is an identity matrix. The fixed parameters (i.e. the constant components of the scale factors) and bias vector components were estimated with a classical least-squares parametric compensation. The remaining parameters were estimated with variance analysis techniques, as explained in [30, 57].

The adaptive correlation-based diagonal dynamic weighting matrix $\mathbf{J}_k$, which is used for identifying the external perturbation using the correlation analysis is defined as

$$\mathbf{J}_k = \begin{bmatrix} \dfrac{\omega_{k,x}}{\|\omega_{k,\mathbf{X}}\|} & 0 & 0 \\[2ex] 0 & \dfrac{\omega_{k,y}}{\|\omega_{k,\mathbf{Y}}\|} & 0 \\[2ex] 0 & 0 & \dfrac{\omega_{k,z}}{\|\omega_{k,\mathbf{Z}}\|} \end{bmatrix}, \tag{13}$$

while the individual component of the dynamic weight vector $\omega_k \in \mathbb{R}^3$ is defined as the correlation between each axis and the corresponding axes that are measuring the same direction of the magnetic field in the virtual measurement point p. For the triple measurement, we can define

$$\omega_k = [\, |\rho(x_\eta, x_1, x_2)| \quad |\rho(y_\eta, y_1, y_2)| \quad |\rho(z_\eta, z_1, z_2)| \,]^T, \tag{14}$$

while the normalization term in the diagonal weighting matrix for the number of n sensors for an arbitrary axis $\mathbf{N} \forall k \in \mathbb{Z},\ k = 1, \ldots, 3$ of the sensor k can be defined from all the corresponding axes using

$$\|\omega_{k,\mathbf{N}}\| = \sum_{i=1}^{n} |\omega_{i,\mathbf{N}}|. \tag{15}$$

Eventually, the joint correlation coefficient ρ between the current axis measurement (e.g. x_η) of the sensor k and the two similar non-orthogonal axis (e.g. x_1 and x_2) can be calculated by

$$\rho(\eta, x_1, x_2) = \sqrt{\frac{r_{\eta,x_1}^2 + r_{\eta,x_2}^2 + 2r_{\eta,x_1} r_{\eta,x_2} r_{x_1,x_2}}{(1 - r_{x_1,x_2}^2) + \epsilon}}, \tag{16}$$

while ϵ is usually taken as a small value to prevent dividing by zero. Based on the basic concept of the Pearson correlation coefficient [58], r_{χ_1,χ_2} between two sensor measurements (i.e. the measurement η and j in Eq. (16)) can be defined as

$$r_{\chi_1,\chi_2} = \frac{\sum_{i=1}^{\lambda_t} (\chi_{1_i} - \overline{\chi}_1)(\chi_{2_i} - \overline{\chi}_2)}{\sqrt{\sum_{i=1}^{\lambda_t} (\chi_{1_i} - \overline{\chi}_1)^2} \sqrt{\sum_{i=1}^{\lambda_t} (\chi_{2_i} - \overline{\chi}_2)^2}}, \tag{17}$$

while the Pearson correlation coefficient is a measure of the linear correlation between two variables χ_1 and χ_2 and has a value between -1 and $+1$, where 1 is a total positive correlation, 0 is no correlation, and -1 is a total negative correlation.

The last λ_t (i.e. $t = 100$ ms at 100 Hz sampling rate system results a queue that contains the last 10 observations) sets of the measured data have been considered for the joint correlation calculation. In other words, a queue with a size of λ_t is used for calculating the average values $\overline{\chi}_1$ and $\overline{\chi}_2$ in the correlation calculation.

4.2. *Training and test data*

We use a comprehensive dataset for training and testing, which we recorded it in different indoor environments using a 3-DoF gimbal system. Our datasets, which we will refer to as the LMT datasets in the following, contain the raw data of the tetrahedron magnetic field and inertial sensor observation associated with the absolute orientations provided by the 3-DoF gimbal's absolute sensors. The recorded dataset covers different conditions for each of the non-disturbing sets, static magnetic distortion, and environment with dynamic magnetic anomalies. We recorded the data in various indoor environments ranging from a general office area to a kitchen and corridor. Each heading estimation set has a length of 60 to 1800 s and varying rotational and translational velocity. The absolute orientation as ground truth (with $0.1°$ resolution) and the raw sensor data are recorded with a frequency of 100 Hz (i.e. $\Delta_t = 10$ ms). The designed sensing platform uses a 9-axis System in Package (SiP) that combines two chips: the MPU-6500, which contains a 3-axis gyroscope and a 3-axis accelerometer; and AK8963C, a 3-axis magnetometer sensor from Asahi Kasei Microdevices Corporation. In total, the recorded dataset contains ≈ 1.2 million data samples ($d = 200$ min).

4.3. *RNN-based sequence learning and fusion*

As shown by the rotation matrices in Eq. (6), and the magnetic filter model in Eq. (9), the trigonometric components make the data relationship inherently nonlinear for the heading estimation. The solution space for the skew angles could therefore not be easily navigated. To this end, we utilized a recurrent neural network as a fusion model for heading estimation. RNNs are neural networks that have memories for temporal learning. RNNs can process sequential data since they leverage historical information for current state prediction and learn dependencies in a sequence. Therefore, RNNs are well suited for time series data and complex nonlinear problems that involve temporal and spatial information [59]. The framework that we used in this paper is based on Long Short-Term Memory (LSTM), one of the most popular types of RNNs, as they can learn long-term dependencies by introducing memory gates and units. The gates on such an LSTM framework automatically decide when to store or discard data in a training phase, which lets them detect the pattern within the data. Figure 8 shows the different LSTM-based architectures we investigated.

We separated the recorded dataset into a training set (≈ 1 Million samples, $d = 170$ min), a validation set ($60\,k$ samples, $d = 10$ min), and a test set ($120\,k$ samples, $d = 20$ min). As input for the network, we take the present acceleration

326   *M. Karimi et al.*
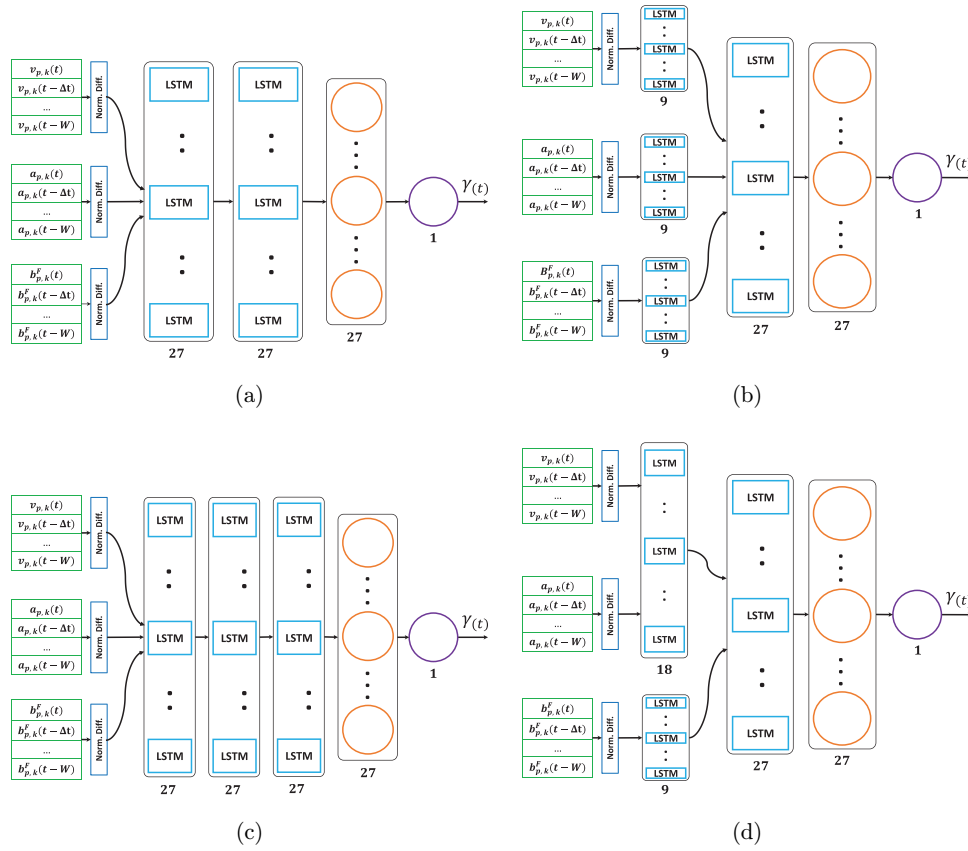
Fig. 8.   Investigated neural network architectures: (a) "LSTM (Simple)", (b) "LSTM (All-Subdivided)", (c) "LSTM (Dense)", (d) "LSTM (Mag-Subdivided)".

$(v_{p,k}(t))$, angular velocity $(a_{p,k}(t))$, and raw or filtered magnetic field $(b_{p,k}(t)/b_{p,k}^{F}(t))$ data as well as past values thereof within a certain time window λ. $\lambda = 0$ ms would imply to consider only the current observation.

The Adam optimizer was used with a learning rate of 1×10^{-4} and with a batch size of 32 samples. A learning rate decay scheduler was used to decrease the initial learning rate of 1×10^{-5} every 10 epochs. The neural networks were trained for 100–500 epochs, depending on the data input selection. After each of the LSTM layers, a dropout layer with a drop rate of 15% is added to circumvent overfitting. The last layers are fully connected, as usual, with 27 and 1 unit, respectively, to provide the relative heading output where $\gamma(t)$ represents the estimated heading angle. We further used the Rectified Linear Unit (ReLU) as an activation function for the fully connected layers. The mean absolute error (MAE) was deployed as a loss function and showed superior performance compared to the mean squared error (MSE). The RNN models were implemented in Keras and Tensorflow. The inference time of the heading angle is in the domain of single-digit millisecond (Dual Core ARM Cortex® A7 @1.6 GHz) and, hence, suitable for real-time applications.

5. Experimental Results

5.1. *Setup*

We developed a skewed redundant sensor platform and used the described fusion model to estimate the heading angle in an indoor environment. We designed a 3D printed frame for mounting the sensors on the structure. Figure 9 shows the sensor
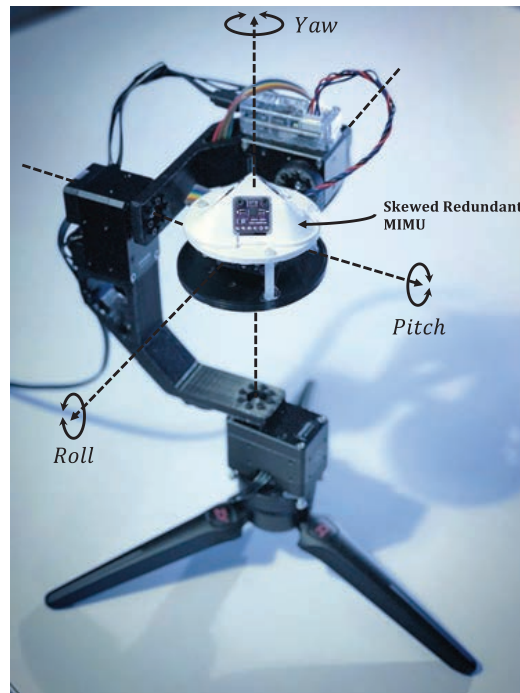
Fig. 9.   Setup of the skewed redundant MIMU platform mounted on a 3-DoF gimbal.

platform and the 3-DoF gimbal used to record the dataset and for system evaluation. This setup has three joints perpendicular to each other, allowing to rotate $360°$ freely in the space. Each of the actuators are equipped with an absolute sensor providing a 12-bit resolution joint pose. As shown in the figure, roll, pitch, and yaw axes can be controlled independently from one another. We placed the platform in different locations using a linear laser ruler to keep the heading angle with respect to the known reference in an indoor environment. We consider the offset of the heading reference that is calculated from the absolute shaft encoder to the Earth's true magnetic heading using the Earth's magnetic model [48]. Different patterns were used to rotate the sensor along each axis. We applied random rotational velocity for each joint. The maximum angular velocity during tests was set to $360°/s$.

To illustrate the gimbal motion during a simple test, Fig. 10 depicts the recorded observations. As shown in this plot, the sensor's final heading is calculated based on the gimbal kinematic, and it is shown as an absolute heading (GT). Angular velocity,

328   *M. Karimi et al.*
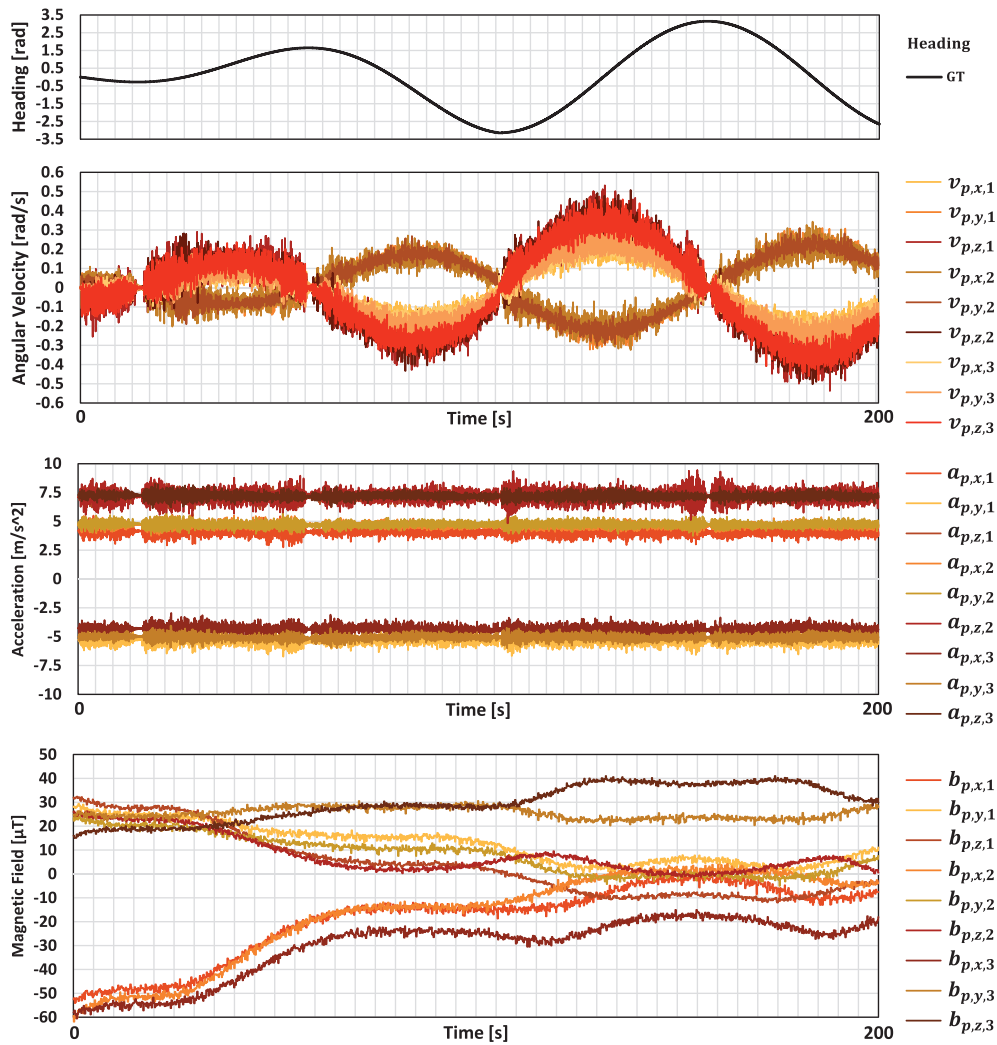
Fig. 10.   Illustration of how the skewed redundant raw data from the sensor observation formed for a motion during a simple test.

acceleration, and the magnetic field in all nine axes are transformed considering the skewed configuration.

5.2. *Performance analysis*

In order to evaluate the proposed method, first, we demonstrate the correlation-based magnetic field filter performance. In [43], it has been shown that saturated channels in multi-axis magnetic field observation can be compensated using the correlation-based filter. However, in this work, we modify this filter to independently perform on each of the 3-axis observations. The output of the filter for the affected observation is shown in Fig. 11. The effect of the perturbation on the raw data $b_{p,y,1}(t)$ can be seen (e.g. around time = 50 s). This data, therefore, is filtered using

Deep Fusion of a Skewed Redundant Magnetic and Inertial Sensor for Heading State 329
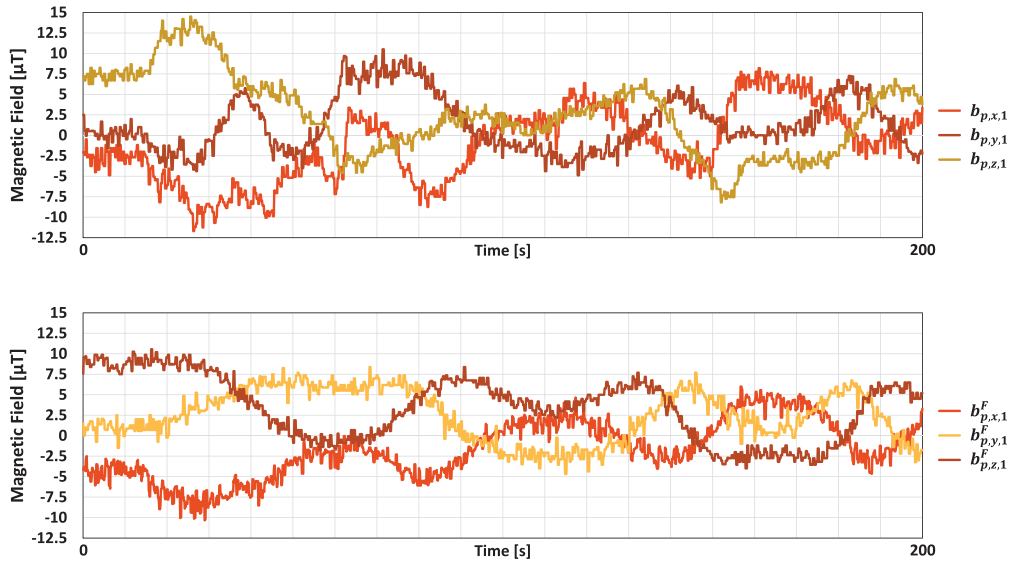
Fig. 11. Result of the correlation-based magnetic field filter for perturbation mitigation using redundant configuration. (Top) Raw data from a 3-axis magnetometer sensor measuring the magnetic field. (Bottom) Filtered data.

the correlation analysis of all the other observed channels. It can be seen that $b^F_{p,y,1}(t)$ is fully eliminated, which shows the improvement and robustness in the presence of the external magnetic field anomaly.

In order to evaluate the heading estimation performance, we use two metrics, mean absolute error (MAE), and root-mean-square error (RMSE). Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. This means the RMSE is most useful when large errors are particularly undesirable. The effect of the window size for the final heading estimation is shown in Table 1. We found out empirically that considering the 70 last values for each heading estimate ($\lambda = 700\,\text{ms}$) delivers the best results. Overall, the LSTM structure with all-subdivided input showed the lowest error with correlation-based magnetic

Table 1. Qualitative comparison of different RNN models by means of the Mean Absolute Error (MAE) and Root-Mean-Square Error (RMSE) for different window size between 0.1–1.0 s.

RNN models [LMT dataset]		All raw data $(\mathbf{a}_{p,k}, \mathbf{v}_{p,k}, \mathbf{b}_{p,k})$					Raw $(\mathbf{a}_{p,k}, \mathbf{v}_{p,k})$ + Filtered $(\mathbf{b}^F_{p,k})$				
Method Window λ [s]		1.0	0.7	0.5	0.3	0.1	1.0	0.7	0.5	0.3	0.1
MAE [°]	LSTM (Dual)	8.41	9.03	8.17	10.51	13.67	4.72	3.91	4.61	6.84	7.17
	LSTM (All-subdivided)	6.83	5.41	7.93	9.49	11.73	**2.28**	**2.11**	**2.93**	**3.74**	**5.62**
	LSTM (Dense)	7.45	8.06	7.80	8.27	12.09	3.76	3.18	**2.98**	3.56	5.89
	LSTM (Mag-subdivided)	7.61	8.12	8.01	10.45	13.36	3.71	3.26	3.01	4.11	6.27
RMSE [°]	LSTM (Dual)	8.90	8.51	9.36	11.29	18.39	4.14	4.02	4.38	5.65	8.68
	LSTM (All-subdivided)	7.91	7.72	9.15	10.60	13.37	**3.70**	**2.48**	**3.08**	**5.18**	**7.44**
	LSTM (Dense)	8.65	9.21	8.55	10.81	18.02	4.09	3.91	**3.41**	4.72	7.28
	LSTM (Mag-subdivided)	9.05	8.96	8.52	10.16	16.24	3.92	3.80	4.16	6.42	8.02

330   *M. Karimi et al.*

field filtering. The proposed neural networks are trained on a subset of the LMT dataset and validated on the complementary set of data. We achieved a RMSE of less than $2.5°$ for heading estimation using the LMT dataset. The second best-performing structure is the LSTM dense model, where the RMSE error increases to $3.0°$ for the window size of $\lambda = 500$.

In Fig. 12, a simple example where the sensor is rotating in three-dimensional space using the gimbal is shown. The absolute heading is available as ground truth
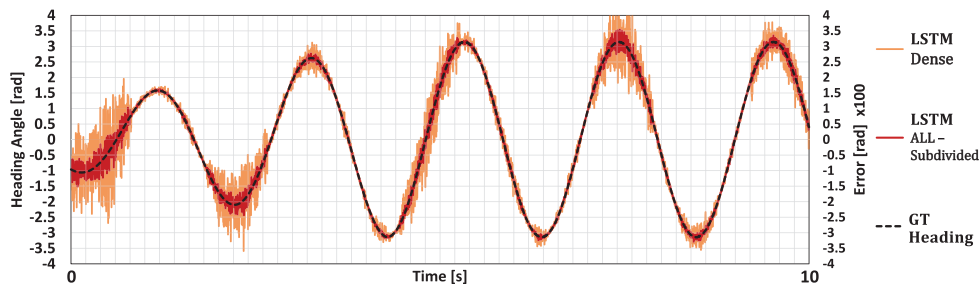
Fig. 12.   Heading estimation error for fast motion using the two best performing models in comparison to the GT heading. The errors are shown with $\times 100$ magnification rate.

(GT), and the error of the two best performing heading estimations are visualized. All the models are highly dependent on the number of available data within the window size. If the data is not available, the accuracy of the estimation degrades dramatically. This issue can be seen at the beginning of the motion, where we feed fewer data to the network, and as soon as the window size reached its maximum amount of observations, the estimation error decreases quickly. In addition, we can see that the estimation error increases when the rotational velocity gets close to zero. This mainly can be interpreted as there was a fast change in the angle, and in the meanwhile, all the gyros present near-zero observations for all the axes. This means the heading estimation was mainly estimated from the magnetic field measurements.

The proposed fusion method has been tested in multiple scenarios portraying different magnetic configurations artificially introduced by permanent magnets or ferromagnetic materials. Additionally, we used two different scenarios for analyzing the system's performance for long-time heading estimation compared to the state-of-the-art methods. For the first test, a stationary platform mounted on a gimbal is used. For the second scenario, while the system experienced the translational motion, the heading estimation is performed. We compared the two best-performing proposed structures with two different fusion models (see Table 2). The Extended Kalman Filter is used with single or with skewed redundant MIMU observations. Additionally, a quaternion-based attitude estimation filter (known as Madgwick filter) with gyroscope-based correction steps and automatic fusion weight adaptation are used for comparison [60]. In long-term heading estimation (e.g. 60 min), the heading estimation error reached $3.2°$. As shown in Table 2, for a short time (e.g. 1.0 min), the estimation error with the proposed models is higher than for the

Chapter 4. Deep Fusion of a Skewed Redundant Magnetic and Inertial Sensor for Heading State
Estimation in a Saturated Indoor Environment

64

Deep Fusion of a Skewed Redundant Magnetic and Inertial Sensor for Heading State 331

Table 2. MAE and RMSE of the proposed fusion system compared with the Kalman filtered approach as well as the Madgwick-based fusion. The measurement is a result of observation for different use time with external magnetic field perturbation in an indoor environment.

Experiments		Rotation only (+Anomaly)				Dynamic motion (+Anomaly)			
	Method time [m]	1.0	15.0	30.0	60.0	1.0	15.0	30.0	60.0
MAE [°]	EKF (Single)	1.02	2.37	3.09	3.44	1.15	2.60	2.93	3.55
	EKF (Skewed redundant)	0.96	2.09	2.72	2.92	**0.93**	2.16	2.89	3.24
	Madgwick (Single)	**0.82**	2.48	3.15	3.46	1.97	2.23	3.18	3.82
	LSTM (Dense)	4.19	1.89	2.30	2.92	5.18	2.21	2.68	3.12
	LSTM (All-subdivided)	4.70	**1.46**	**1.88**	**1.91**	4.48	**1.82**	**2.11**	**2.91**
RMSE [°]	EKF (Single)	2.03	2.85	3.58	3.97	2.73	3.01	3.82	4.27
	EKF (Skewed redundant)	1.82	**1.98**	3.02	3.16	2.06	**2.19**	3.72	3.91
	Madgwick (Single)	**1.67**	3.14	3.97	4.41	**1.91**	3.57	4.16	5.26
	LSTM (Dense)	5.12	1.82	2.59	3.09	5.60	2.98	3.12	3.62
	LSTM (All-subdivided)	5.73	2.07	**1.96**	**2.18**	5.92	2.39	**2.47**	**3.21**

conventional methods. This is mainly due to the need for an observation of at least 500 ms of window size.

6. Conclusions

This paper proposed a skewed redundant magnetometer and inertial sensor fusion approach for robust heading estimation in an indoor environment. We exhibited heading estimation using the Earth's magnetic field and the different perturbation sources for heading angle estimation in the presence of the external magnetic anomalies. We showed the Hall-effect magnetic sensors' principle and analyzed the external magnetic field's source on 3-axis MEMS-based sensors. Inspired by the skewed redundant configuration, we designed a tetrahedron platform using a triple 9-axis sensor to detect and mitigate the saturated channels on the observation signals. In order to pre-filter the sensory data, we proposed a correlation-based filter model for multi-magnetometer platforms. RNN-based sensor fusion models are presented and used for skewed redundant MIMU fusion. We investigated multiple structures using LSTM models. We further investigated the models' performance in different indoor environments for different scenarios and validated the algorithm using the internally recorded dataset. Additionally, experiments using a gimbal platform have been conducted in diverse scenarios to verify the generality of the proposed fusion method by means of robust heading estimation using the skewed redundant magnetometer platform. Experiments showed that although the proposed system is reliable for heading estimation in long-term use, it has a high dependency on the input window size. As a result, for short-term heading estimation, the performance of the conventional methods was similar or better than the proposed model. However, due to the observed drifts in long-term use, the proposed system provides a more reliable estimation with less deviation in saturated indoor environment.

332   *M. Karimi et al.*

In future work, we plan to investigate on more complex deep neural networks to enhance the quality of the estimation. Additionally, in cases where there is a constructive disturbance along with changes in user dynamics, the perturbation mitigation approach will be further enhanced to take into account the uncertainty between perturbation and agent motion. Furthermore, the effect of the number of the sensing axes on the estimation of the true Earth's magnetic field needs to be studied using analytical models.

Acknowledgments

References

[1] R. Harle, A survey of indoor inertial positioning systems for pedestrians, *IEEE Commun. Surv. Tut.* **15**(3) (2013) 1281–1293.

[2] O. S. Salychev, *Applied Inertial Navigation: Problems and Solutions* (BMSTU Press Moscow, Russia, 2004).

[3] T. Aykut, M. Karimi, C. Burgmair, A. Finkenzeller, C. Bachhuber and E. Steinbach, Delay compensation for a telepresence system with 3D 360° vision based on deep head motion prediction and dynamic FoV adaptation, *IEEE Robot. Automat. Lett.* **3**(4) (2018) 4343–4350.

[4] T. Aykut, C. Burgmair, M. Karimi, J. Xu and E. Steinbach, Delay compensation for actuated stereoscopic 360 degree telepresence systems with probabilistic head motion prediction, in *2018 IEEE Winter Conf. Applications of Computer Vision*, 2018, pp. 2010–2018.

[5] P. H. Mathisen and T. I. Fossen, Robust navigation system for UAVs in GNSS-and magnetometer-denied environments, in *2019 Int. Conf. Unmanned Aircraft Systems*, 2019, pp. 1416–1424.

[6] T. Bailey, J. Nieto, J. Guivant, M. Stevens and E. Nebot, Consistency of the EKF-SLAM algorithm, in *2006 IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2006, pp. 3562–3568.

[7] J. Kelly and G. S. Sukhatme, Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration, *Int. J. Robot. Res.* **30**(1) (2011) 56–79.

[8] M. Karimi, S. S. Ghidary, R. Shekhar, T. D. Kane and R. Monfaredi, Magnetically anchored pan-tilt stereoscopic robot with optical-inertial stabilization for minimally invasive surgery, in *Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*, Vol. 10951 (International Society for Optics and Photonics, 2019), p. 109511E.

[9] J. Zhang and S. Singh, Visual-lidar odometry and mapping: Low-drift, robust, and fast, in *2015 IEEE Int. Conf. Robotics and Automation*, 2015, pp. 2174–2181.

[10] R. Mur-Artal and J. D. Tardós, Visual-inertial monocular SLAM with map reuse, *IEEE Robot. Automat. Lett.* **2**(2) (2017) 796–803.

[11] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart and P. Furgale, Keyframe-based visual–inertial odometry using nonlinear optimization, *Int. J. Robot. Res.* **34**(3) (2015) 314–334.

[12] K. Gade, The seven ways to find heading, *J. Navigat.* **69**(5) (2016) 955–970.

[13] J. Wu, Z. Zhou, J. Chen, H. Fourati and R. Li, Fast complementary filter for attitude estimation using low-cost MARG sensors, *IEEE Sensors J.* **16**(18) (2016) 6997–7007.

[14] M. Kok and T. B. Schön, Magnetometer calibration using inertial sensors, *IEEE Sensors J.* **16**(14) (2016) 5679–5689.

[15] J. F. Vasconcelos, G. Elkaim, C. Silvestre, P. Oliveira and B. Cardeira, Geometric approach to strapdown magnetometer calibration in sensor frame, *IEEE Trans. Aerospace Electron. Syst.* **47**(2) (2011) 1293–1306.

[16] M. Kok, J. D. Hol, T. B. Schön, F. Gustafsson and H. Luinge, Calibration of a magnetometer in combination with inertial sensors, in *2012 15th Int. Conf. Information Fusion*, 2012, pp. 787–793.

[17] J. Lenz and S. Edelstein, Magnetic sensors and their applications, *IEEE Sensors J.* **6**(3) (2006) 631–649.

[18] M. Kok and T. B. Schön, Maximum likelihood calibration of a magnetometer using inertial sensors, *IFAC Proc. Vol.* **47**(3) (2014) 92–97.

[19] K. Han, H. Han, Z. Wang and F. Xu, Extended Kalman filter-based gyroscope-aided magnetometer calibration for consumer electronic devices, *IEEE Sens. J.* **17**(1) (2016) 63–71.

[20] X. Tu, Magnetometer calibration, US Patent 8,717,009 (2014).

[21] P. L. Worcester, D. J. Maxwell and M. L. Trethewey, Calibration of magnetic compasses, US Patent 5,117,375, (1992).

[22] J. Fang, H. Sun, J. Cao, X. Zhang and Y. Tao, A novel calibration method of magnetic compass based on ellipsoid fitting, *IEEE Trans. Instrument. Measurement* **60**(6) (2011) 2053–2061.

[23] M. Zhu, W. Ouyang and Y. Wu, Orientation estimation by partial-state updating Kalman filter and vectorial magnetic interference detection, *IEEE Trans. Aerospace Electron. Syst.* (2021).

[24] I. Skog, Inertial and magnetic-field sensor arrays-capabilities and challenges, in *2018 IEEE Sensors*, 2018, pp. 1–4.

[25] M. H. Afzal, V. Renaudin and G. Lachapelle, Assessment of indoor magnetic field anomalies using multiple magnetometers, in *The Institute of Navigation Global Navigation Satellite System*, Vol. 10, 2010, pp. 21–24.

[26] V. Renaudin, M. H. Afzal and G. Lachapelle, Complete triaxis magnetometer calibration in the magnetic domain, *J. Sensors* **2010** (2010).

[27] M. H. Afzal, V. Renaudin and G. Lachapelle, Use of earth's magnetic field for mitigating gyroscope errors regardless of magnetic perturbation, *Sensors* **11**(12) (2011) 11390–11414.

[28] V. Renaudin, M. H. Afzal and G. Lachapelle, New method for magnetometers based orientation estimation, in *IEEE/ION Position, Location and Navigation Symp.*, 2010, pp. 348–356.

[29] G. Yin and L. Zhang, Magnetic heading compensation method based on magnetic interferential signal inversion, *Sens. Actuat. A: Phys.* **275** (2018) 1–10.

[30] S. Guerrier, Improving accuracy with multiple sensors: Study of redundant MEMS-IMU/GPS configurations, in *Proc. 22nd Int. Technical Meeting of the Satellite Division of the Institute of Navigation*, 2009, pp. 3114–3121.

334   *M. Karimi et al.*

[31] X. Zhang, L. Xiao, C. Zou, L. Guo, Y. Jin and G. Shi, A novel calibration method of electronic compass based on multi-redundancy, in *2017 IEEE Int. Conf. Real-time Computing and Robotics*, 2017, pp. 156–161.

[32] I. Colomina, M. Giménez, J. Rosales, M. Wis, A. Gómez and P. Miguelsanz, Redundant IMUs for precise trajectory determination, Vol. 1223, 2004, p. 17.

[33] S. Sukkarieh, P. Gibbens, B. Grocholsky, K. Willis and H. F. Durrant-Whyte, A low-cost, redundant inertial measurement unit for unmanned air vehicles, *Int. J. Robot. Res.* **19**(11) (2000) 1089–1103.

[34] R. Giroux, S. Sukkarieh and M. Bryson, Implementation of a skewed-redundant low-cost ins in a fast-prototyping environment, in *Proc. Institute of Navigation National Technical Meeting*, 2004.

[35] M. Jafari, Optimal redundant sensor configuration for accuracy increasing in space inertial navigation system, *Aerospace Sci. Technol.* **47** (2015) 467–472.

[36] A. Osman, B. Wright, A. Noureldin and N. El-Sheimy, Multi-sensor inertial navigation systems employing skewed redundant inertial sensors, Int. Navigation Global Navigation Satellite System — 19th Int. Technical Meeting of the Satellite Division, 2006, pp. 26–29.

[37] Y. Yu and M. Cheng, Comparative research of redundant strap down inertial navigation system based on different configuration schemes, in *American Institute Physics Conf. Proc.*, Vol. 1967, 2018, p. 040026.

[38] M. H. Afzal, V. Renaudin and G. Lachapelle, Multi-magnetometer based perturbation mitigation for indoor orientation estimation, *Navigation* **58**(4) (2011) 279–292.

[39] H. Naseri and M. Homaeinezhad, Improving measurement quality of a MEMS-based gyro-free inertial navigation system, *Sens. Actuat. A Phys.* **207** (2014) 10–19.

[40] S. Ghasemi-Moghadam and M. Homaeinezhad, Attitude determination by combining arrays of MEMS accelerometers, gyros, and magnetometers via quaternion-based complementary filter, *Int. J. Num. Mod. Electron. Netw. Dev. Fields* **31**(3) (2018) e2282.

[41] M. W. Givens and C. Coopmans, A survey of inertial sensor fusion: Applications in sUAS navigation and data collection, in *2019 Int. Conf. Unmanned Aircraft Systems*, 2019, pp. 1054–1060.

[42] F. Liu, Z. Su, H. Zhao, Q. Li and C. Li, Attitude measurement for high-spinning projectile with a hollow MEMS IMU consisting of multiple accelerometers and gyros, *Sensors* **19**(8) (2019) 1799.

[43] M. Karimi, E. Babaians, M. Oelsch, T. Aykut and E. Steinbach, Skewed-redundant hall-effect magnetic sensor fusion for perturbation-free indoor heading estimation, in *2020 Fourth IEEE Int. Conf. Robotic Computing*, 2020, pp. 367–374.

[44] A. Grosz, M. J. Haji-Sheikh and S. C. Mukhopadhyay, *High Sensitivity Magnetometers* (Springer, 2017).

[45] R. Hansenw, R. L. Kleinberg and A. Kaufman, *Principles of the Magnetic Methods in Geophysics*, Methods in Geochemistry and Geophysics, Vol. 42 (Elsevier, 2009).

[46] E. Ramsden, *Hall-effect Sensors: Theory and Application* (Elsevier, 2011).

[47] S. Tumanski, *Handbook of Magnetic Measurements* (CRC Press, 2016).

[48] Global geomagnetic models (2019).

[49] M. Sturza, Skewed axis inertial sensor geometry for optimal performance, in *Digital Avionics Systems Conf.*, 1988, p. 3874.

[50] S. Qiu, Z. Wang, H. Zhao, K. Qin, Z. Li and H. Hu, Inertial/magnetic sensors based pedestrian dead reckoning by means of multi-sensor fusion, *Inf. Fusion* **39** (2018) 108–119.

[51] H. Liu, R. Shor and S. S. Park, Intelligent filter for accurate subsurface heading estimation using multiple integrated MEMS sensors, in *2018 IEEE Sensors*, 2018, pp. 1–4.

[52] H. Hou, *Modeling Inertial Sensors Errors Using Allan Variance* (University of Calgary, 2004).

[53] Z. Xing and D. Gebre-Egziabher, Modeling and bounding low cost inertial sensor errors, in *2008 IEEE/ION Position, Location and Navigation Symp.*, 2008, pp. 1122–1132.

[54] G. Troni and R. M. Eustice, Magnetometer bias calibration based on relative angular position: Theory and experimental comparative evaluation, in *2014 IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2014, pp. 444–450.

[55] B. Dil, G. Hendeby, F. Gustafsson and B. J. Hoenders, Approximate diagonalized covariance matrix for signals with correlated noise, in *2016 19th Int. Conf. Information Fusion*, 2016, pp. 521–527.

[56] D. Unsal and K. Demirbas, Estimation of deterministic and stochastic IMU error parameters, in *Proc. 2012 IEEE/ION Position, Location and Navigation Symp.*, 2012, pp. 862–868.

[57] Y. Stebler, S. Guerrier and J. Skaloud, An approach for observing and modeling errors in MEMS-based inertial sensors under vehicle dynamic, *IEEE Trans. Instrument. Measurement* **64**(11) (2015) 2926–2936.

[58] J. Benesty, J. Chen, Y. Huang and I. Cohen, Pearson correlation coefficient, in *Noise Reduction in Speech Processing* (Springer, 2009), pp. 1–4.

[59] D. Weber, C. Gühmann and T. Seel, Neural networks versus conventional filters for inertial-sensor-based attitude estimation, in *2020 IEEE 23rd Int. Conf. Information Fusion*, 2020, pp. 1–8.

[60] S. O. Madgwick, A. J. Harrison and R. Vaidyanathan, Estimation of IMU and MARG orientation using a gradient descent algorithm, in *2011 IEEE Int. Conf. Rehabilitation Robotics*, 2011, pp. 1–7.

Chapter 5

LoLa-SLAM: Low-Latency LiDAR SLAM Using Continuous Scan Slicing

5.1 Low-Latency 6D LiDAR Localization and Mapping for Real-Time Indoor Applications

The utilization of real-time 6D pose estimation is a key component for autonomous indoor navigation of mobile robots, specially for Unmanned Aerial Vehicle (UAV)s. In this chapter we present a low-latency SLAM framework based on LiDAR scan slicing and concurrent matching, called LoLa-SLAM. The proposed framework uses sliced point cloud data from a rotating LiDAR in a concurrent multi-threaded matching pipeline. This way the 6D ego-motion estimation with high update rate and low latency is guaranteed. The LiDAR sensor is actuated by a 2-DoF Lissajous spinning pattern, which overcomes the sensor's limited field of view. The present work proposes to estimate the two-dimensional roughness of a point cloud in order to extract feature points for fine matching and registration of the set of points. Furthermore, the pose estimator employs a temporal motion predictor to assist in finding the feature correspondences in the domain map, which aids the nonlinear optimizer in achieving a fast convergence rate. Taking this as a second step, a Extended Kalman Filter (EKF) is used for the final ego-motion estimation. Multiple experiments were performed to evaluate the framework by comparing the accuracy, the latency, and the update rate of the pose estimation for the trajectory flown in an indoor environment. Based on the comparison between the state-of-the-art frameworks and the generated volumetric map, we quantify the superior quality of the generated voxel grid map. Additionally, measured ground truth position information recorded from a total station unit was used to examine the precision of ego-motion estimation.

5.2   Authors Contribution

Conceptualization, design, and development of the measurement method using actuated LiDAR platform, the UAV system, the actuation technique, point cloud slicing solution for low-latency ego-motion estimation, and execution of the experiments were performed by Mojtaba Karimi. Data visualization were partially conducted by Oliver Stengel and Edwin Babaians. Recording the data was conceptualized by Mojtaba Karimi and Martin Oelsch. Mojtaba Karimi and Oliver Stengel evaluated the experimental results. Eckehard Steinbach supervised this work. The original draft was written by Mojtaba Karimi. The paper was reviewed and edited by all authors.

Title:

LoLa-SLAM: Low-Latency LiDAR SLAM Using Continuous Scan Slicing

Authors:

Mojtaba Karimi, Martin Oelsch, Oliver Stengel, Edwin Babaians, and Eckehard Steinbach

LoLa-SLAM: Low-Latency LiDAR SLAM Using Continuous Scan Slicing

Mojtaba Karimi [ID], *Member, IEEE*, Martin Oelsch [ID], *Member, IEEE*, Oliver Stengel, Edwin Babaians [ID], and Eckehard Steinbach [ID], *Fellow, IEEE*

Abstract—**Real-time 6D pose estimation is a key component for autonomous indoor navigation of Unmanned Aerial Vehicles (UAVs). This letter presents a low-latency LiDAR SLAM framework based on LiDAR scan slicing and concurrent matching, called LoLa-SLAM. Our framework uses sliced point cloud data from a rotating LiDAR in a concurrent multi-threaded matching pipeline for 6D pose estimation with high update rate and low latency. The LiDAR is actuated using a 2D Lissajous spinning pattern to overcome the sensor's limited FoV. We propose a two-dimensional roughness model to extract the feature points for fine matching and registration of the point cloud. In addition, the pose estimator engages a temporal motion predictor that assists in finding the feature correspondences in the map for the fast convergence of the non-linear optimizer. Subsequently, an Extended Kalman Filter (EKF) is adopted for final pose fusion. The framework is evaluated in multiple experiments by comparing the accuracy, latency, and the update rate of the pose estimation for the trajectories flown in an indoor environment. We quantify the superior quality of the generated volumetric map in comparison to the state-of-the-art frameworks. We further examine the localization precision using ground truth pose information recorded by a total station unit.**

Index Terms—**SLAM, aerial systems, perception and autonomy, low-latency localization.**

I. INTRODUCTION

THE demand for unmanned aerial vehicles (UAVs) for autonomous exploration and inspection is growing. The compact design-factor, relatively low cost, and maneuverability of UAVs make them well suited for various tasks [1]. Autonomous UAVs require a reliable navigation system to operate in challenging environments, such as GPS-denied or cluttered indoor areas [2], [3]. Tackling the general navigation mission, one needs to address a set of problems ranging from 6D pose estimation to trajectory planning [4], [5]. In practice, the navigation task's performance depends mainly on the accuracy,

Fig. 1. Real-time localization of a UAV in an indoor environment using the proposed SLAM framework. It provides low-latency 6D pose estimation with a high temporal update rate for autonomous navigation while generating a full-scale 3D map of the environment.

update rate, and latency of the perception unit (see Fig. 1). For instance, the trajectory tracking controller of the UAV requires a high-rate pose estimation (in the range of 0.1 to 1.0 kHz) with comparatively low latency (maximum 50 ms) to achieve a fully autonomous flight [6]–[8]. In this context, latency is the time difference from the sensor observation until the localization algorithm provides the estimated pose.

Visual-inertial localization and mapping is well studied for drone localization as it can meet the aforementioned requirements [9], [10]. However, in large indoor areas, the accuracy of visual-inertial localization approaches tends to degrade dramatically [11], [12]. This is mainly due to the considerable distance between the camera and the scene, which causes feature tracking failures in the camera frames. As an alternative, laser scanning technology is employed to capture precise range measurements enabling LiDAR-based localization and mapping systems. To address the 6D pose estimation task, three-dimensional simultaneous localization and mapping (3D SLAM) is the preferred way. 3D SLAM is mainly addressed in the literature by using multiple laser scanners, which are installed in both horizontal and vertical frames [13], [14]. However, due to the weight and processing constraints in UAVs, in practice, the use of an actuated LiDAR is preferred rather than installing multiple or heavier sensors with an inherently larger FoV [15], [16]. In this regard, different spinning mechanisms have been developed to be mounted on drones [17], [18]. Although generally successful, existing solutions come with some shortcomings, such as long revisit time, blind spots, inconsistencies in scanning the environment, and skewing issues. These shortcomings cause problems

in determining the correspondences within the point cloud data for 6D positioning [19], [20].

While LiDAR-based SLAM has been studied widely in the past decade for mobile robots [19], [21]–[24], there is still a major gap in utilizing these systems for real-time navigation of UAVs. This is mainly because it is not possible to rely on an aerial vehicle remaining sufficiently motionless. The drone position must be represented at a high temporal update rate with low-latency to achieve stable closed-loop control. Addressing the challenge mentioned above, continuous-time trajectory estimators were developed in the literature [25], [26]. However, due to the limited FoV and the substantial time needed for a complete scan period of the LiDAR, estimating continuous odometry and global registration introduces a significant latency and low-fidelity 6D pose estimation in such a system. To address this issue, LiDAR-inertial SLAM systems use an additional Inertial Measurement Unit (IMU) to produce high-rate pose updates [24]. They provide acceptable results for ground-based mobile robots. However, as UAVs experience high dynamic motions and inevitable high-frequency vibrations due to the propellers' rotation, IMU-based approaches leads to having a fluctuating pose estimation in such applications [27].

To meet these challenges, we propose a low-latency localization and mapping framework using scan slicing (LoLa-SLAM). In our approach, the point cloud data from a rotating laser scanner with 360° horizontal FoV is continuously sliced and used for estimating the real-time 6D pose. Besides, we introduce a novel sensor payload design in which a small, lightweight multi-line 360° laser scanner is actuated with a Lissajous pattern. This structure is employed to reduce the revisit time and provides a dense point cloud with consistency in scanning the environment [28], [29]. To this end, while UAV localization using actuated LiDARs has been presented previously [25], [30]–[32], to the best of our knowledge, this is the first work on scan slicing of an actuated rotating LiDAR, which is utilized for real-time low-latency indoor localization. This work has three main contributions. First, it presents an actuated LiDAR sensor platform's design and implementation based on a centralized servo mechanism and a 2D Lissajous pattern to improve the FoV and reduce the revisit time. Second, this letter proposes a novel real-time LiDAR-based low-latency SLAM framework based on scan slicing and a concurrent multi-threaded matching pipeline. Third, multiple experiments were designed and conducted to evaluate the proposed framework's performance toward the reliable localization of a UAV in an indoor environment.

The rest of this letter is organized as follows: Section II describes the problem and notation used in this work. Section III introduces the sensor platform and the Lissajous LiDAR actuation pattern, and Section IV details the continuous sliced scanning model. The low-latency localization framework is introduced in Section V. Experimental evaluations are presented in Section VI. Finally, Section VII concludes this letter and examines potential future works.

II. PROBLEM STATEMENT AND SENSOR PLATFORM

Rotating laser scanners use the time-of-flight (ToF) concept that operates by emitting laser light and capturing its reflection
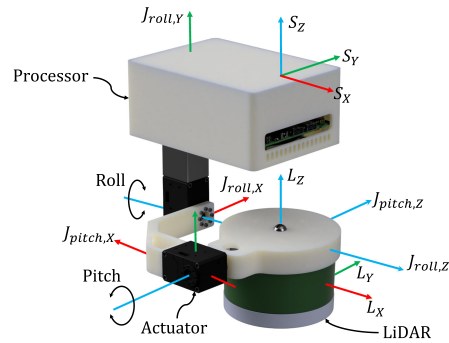
Fig. 2. CAD model of the sensor payload design for UAVs represented with the coordinate frames. A 2-DoF centralized servo mechanism is utilized for LiDAR actuation. The system includes a rotating 16 beam-line LiDAR, 2-DoF centralized servo mechanism for LiDAR actuation, motor drivers, and a processing unit.

to measure distances to nearby objects. In the case of multi-line scanning, multiple laser beams are used with a rotating structure at a fixed velocity to provide a more extended FoV compared to single-line planar LiDARs. Each of the ToF measurements is associated with the horizontal and vertical angle of the corresponding beam. This information is combined to compute the 3D coordinates of the sensed point, and accordingly, a LiDAR generates a point cloud corresponding to the set of observed points. The LoLa-SLAM system studied in this work is validated on the actively actuated LiDAR system shown in Fig. 2. The Velodyne VLP-16 laser scanner has a vertical FoV of 30° and a horizontal FoV of 360° with approximately 0.4° angular resolution and 16 laser beams. The laser scanner is connected to a centralized hinge joint with 2-DoF. Each of the joints can be controlled individually with a maximum velocity of 360°/s in a range of ±45° with 0.087° resolution.

We use right-handed coordinate systems with an uppercase subscript. All laser points are measured in the LiDAR coordinate frame $\{L\}$. As depicted in Fig. 2, the origin of this frame is in the geometric center of the laser scanner. The collected points that belong to one complete rotation of a multi-beam laser in 360° are represented as a single full scan. Classically, each full scan is used in the SLAM algorithms to determine the ego-motion of the platform. In this work, instead of using a full scan of a rotating LiDAR, we use continuously sliced scans as a set of points $\mathcal{P}$ with right subscript q, $q \in Z^+$ to indicate the slice number. Each slice is a subset of an upcoming full scan. Each point in the sliced point cloud $\mathcal{P}$ is defined as i, $i \in \mathcal{P}_q$ in the LiDAR frame $\{L\}$ and is marked by $\mathcal{P}^L_{(q,i)}$. Similarly, the sensor coordinate frame $\{S\}$ is a 3D coordinate system coinciding with $\{L\}$ and the joint frame $\{J_{pitch}\}$ and $\{J_{roll}\}$. Map coordinate frame $\{M\}$ is a coordinate system coinciding $\{S\}$ with the initial position $\{0\}$. A point in the sliced point cloud $\mathcal{P}_q$, in the map coordinate is denoted as $\mathcal{P}^M_{(q,i)}$. With the specified notation defined above, the low-latency localization and mapping is defined as: given a sequence of continuous sliced scans in the sensor frame $\mathcal{P}^M_q$, estimate the 6D pose of the platform in the map coordinate frame $\{M\}$ as $^M\boldsymbol{T}_{S,q}(t)$, and create a representative map of the environment.

III. LiDAR Actuation Using Lissajous Pattern

The limited vertical FoV of state-of-the-art laser scanners is one of the main challenges for 6D positioning. In robotic applications, this issue is normally addressed by an articulated sensor, where it is driven by one or more dedicated actuators in order to increase the FoV. However, such observation requires a non-negligible amount of time to capture. This becomes problematic for drones while they are floating in the air. Periodic actuation of the LiDAR sensor in roll and/or pitch direction repeats the observation of the surface patches and allows for a larger FoV. To enhance this procedure, researchers have developed non-raster scan actuation motions, called Lissajous pattern [28], [29], [33]. Considering the fact that a state-of-the-art rotating LiDAR provides a constant sampling rate, the main benefit of using a Lissajous scan pattern-based actuation of the LiDAR is the reduced revisit time (by 54%) within the defined FoV [29]. In addition, compared to the other actuation models, avoiding the orderly nature of the raster scan by using a simultaneously periodic sinusoidal trajectory can be defined as another advantage of the Lissajous pattern.

Motivated by these works, we propose a 2D Lissajous pattern for a rotating 360° LiDAR. This pattern allows us to rotate the LiDAR to observe the vertical and horizontal surfaces equitably. We designed and developed an actuated LiDAR payload with 2-DoF, as shown in Fig. 2, to be able to precisely execute this pattern. Based on the systematic study of the scan skewing problem [34], we developed a mechanism that actuates the laser scanner in a combined roll and pitch motion around a centralized axes to minimize the measurement distortion. The platform is manufactured using printed ABS material. The proposed 2D Lissajous scan pattern at time t for the roll (α) and pitch (β) axes are defined by

$$\begin{aligned} \alpha(t) &= A_{roll}(t)\sin(2\pi f_c t), \\ \beta(t) &= A_{pitch}(t)\cos(2\pi f_c t), \end{aligned} \tag{1}$$

where the frequency f_c is the period of one full sweep, and the amplitudes A_{roll} and A_{pitch} are defined as the maximum angles of the motion. We set the period of one full sweep to 8.0 seconds and the maximum amplitude to 45°. These values are obtained heuristically based on the dynamics of the drone and the FoV of the sensor. We use a slow start mechanism, which increases the amplitude of the Lissajous actuation pattern from zero to its maximum value within the first three sweeps. We use the smooth start mainly to prevent damage to the rotating laser scanner as the sudden and fast motion will cause hardware damage in long term use. Although it is not necessary, this smooth increase of the amplitude at the start, while the robot is stationary (e.g., before takeoff), helps in the initialization of the SLAM algorithm. This is because, at the start, there is a large number of the newly observed points, which must be added to the map. The generated Lissajous pattern in comparison with a single axis raster scan is illustrated in Fig. 3.

IV. Continuous Scan Slicing

The laser scanner used in this work rotates with a maximum of 1200 rounds per minute (rpm) and provides 300 000 points per
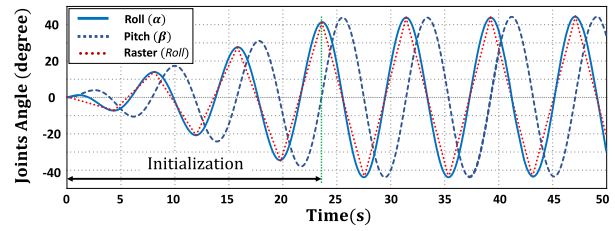
Fig. 3.   The proposed 2D Lissajous pattern for the roll and pitch joints of the sensor platform. The frequency of the pattern is defined to perform one sweep in 8.0 seconds. In the start, a smooth increase in the amplitude is used to minimize potential hardware damage and also to generate the initial map.
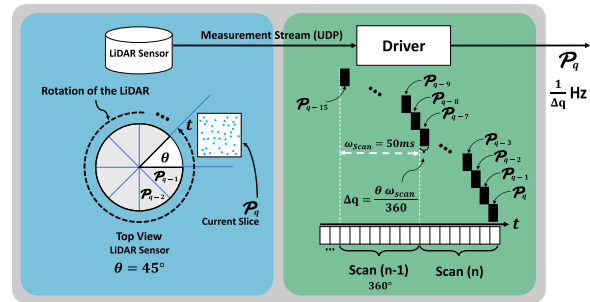
Fig. 4.   Continuous Slicing Point Cloud (CSPC) model. This model gathers smaller collections of measurements that can be used in a concurrent pipeline instead of the normally used full 360° scan. Each slice $\mathcal{P}_q$ is defined for the last $\theta°$ of the rotation of the LiDAR.

second. A 360° scan is available every $\omega_{scan} = 50$ milliseconds and contains 15 thousand measurements. The laser scanner is continuously measuring and creates a constant output stream associated with time and angle. These measurements are accumulated over time at the driver level. When the laser scanner completes a full 360° scan, it releases all the recorded points as a single scan. State-of-the-art SLAM frameworks are developed based on these 360° scans for ego-motion estimation [22], [23], [27]. In practice, the point cloud arrival rate defines the frequency of the pose estimation. However, the ego-motion estimation can be performed after any other measurement, which means 360° is an arbitrary margin, and individual measurements can be bundled freely.

Although it is reasonable to compare complete 360° scans in the odometry unit; we show later in the proposed localization framework that our SLAM model does not depend on the odometry unit. The proposed SLAM framework uses continuous slices of a full 360° scan for low-latency localization using direct scan matching on the global map. The main intention of the proposed continuous slicing point cloud (CSPC) is to gather smaller collections of measurements that can be used in a successive manner instead of waiting for a full 360° scan. In this context, describing the term continuous in other words, we collect the measurements from the upcoming scan into a small point cloud slice for the last θ angle (i.e., only the data of the measurements within the last $\theta = 45°$ of the rotation), and publish them one after another. As shown in Fig. 4, slicing of the points is developed in the driver level of the laser scanner, and the CSPC is published at a rate of $\frac{1}{\Delta q}$ Hz. Considering each slice
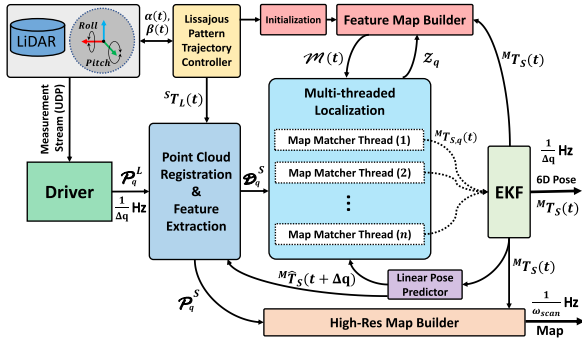
Fig. 5. Overview of the LoLa-SLAM framework. The results of the multi-threaded localization pipeline are fused using the EKF. The feature map is a voxelized map of the collected feature points with respect to the corresponding pose at time t. The voxelized high-resolution map is generated directly from the input CSPC.

is generated from the last $\theta = 45°$ of the rotation of the LiDAR, each slice is available every $\Delta q = 6.25$ ms and noted by $\mathcal{P}_q$ with q indicating the last observed slice. With less angular coverage, the point cloud becomes more ambiguous because opposing features are missing to prevent misguided shifts during optimization in the pose estimation. However, the new measurements can be integrated with old ones to achieve the necessary number of measurements to ensure correct matching. For instance, we can add the new slice ($\theta = 45°$) to the previously registered slices within the last $315°$, in order to obtain a repeatedly full $360°$ scan, however, at a much higher update rate. While the reuse of past measurements does not affect the achieved update rate, the system's real-time performance needs to be carefully investigated due to the increase in computational usage.

V. MULTI-THREADED LOW-LATENCY LOCALIZATION

A. System Overview

The overall framework of the proposed system is depicted in Fig. 5. Let's consider $\mathcal{P}_q^L$ to be the points in the LiDAR coordinate frame $\{L\}$ published in real-time with the update rate of $\frac{1}{\Delta q}$ Hz and the LiDAR to follow the motion with the generated Lissajous pattern. The point cloud $\mathcal{P}_q^L$ is first registered in the sensor coordinate frame using the transformation provided by the absolute status of the joints. During the start phase, the generated map points at time t are denoted with $\mathcal{M}(t)$, and as we assume the platform's pose is in the stationary state; we only collect the extracted feature points. The feature extraction node takes the point cloud $\mathcal{P}_q^S$ in the sensor coordinate frame and extracts the feature points. Let us assume for the first iteration that the minimum initial map is already available from the initialization step.

The multi-threaded localization pipeline takes the previously registered feature points $\mathcal{D}_{q-1}^M$ in combination with the extracted feature points from the current slice $\tilde{\mathcal{D}}_q$, for registration in the map. While, in principle, the current slice could be combined with more than one previous slice, for complexity and latency reasons, we limit it in our implementation to one previous scan slice. The combined feature points $\mathcal{D}_q^S$ are used to compute the

precise motion of the LiDAR from matching the last sweeps within the map. With the slicing process described previously, a single thread is unable to handle the computation in real-time. The computation time increases because the diminished point clouds do not result in equally reduced optimization time, and the reuse of measurements adds additional calculations. Thus, we run the pose estimation in multiple threads with a modern multi-core processor and then fuse the individual concurrent estimates. Each thread extracts the pose transform concerning the map and uses the initial condition of the predicted 6D pose from the linear predictor unit. The Extended Kalman Filter (EKF) unit takes all the concurrent transformations published by the localization unit and fuses them for estimating the final 6D pose. Finally, the map builder unit takes and adds the new feature points into the map with respect to the associated final pose from the EKF unit. Similar to [22], considering that our approach uses a voxelized feature map for pose estimation, and as this unit runs in a dedicated process, a growing map does not affect the real-time localization procedure.

B. Point Cloud Registration and Feature Extraction

The measured point cloud is registered in the sensor coordinate frame using the transformation $^S\boldsymbol{T}_L(t)$. This transformation is obtained by considering the kinematics of the LiDAR actuation. As shown in Fig. 2, the laser scanner is attached to the sensor frame with two actively controlled joints to follow the pattern generated by Eq. (1). The homogeneous transformation is performed for all points i in $\mathcal{P}_{(q,i)}^L$. Let us define the homogeneous 3D affine transformation matrix $\boldsymbol{T}(t) = (\boldsymbol{R}(t), \boldsymbol{t}(t)) \in SE(3)$ as a rigid body transform composed of a rotation matrix $\boldsymbol{R}(t) \in SO(3)$ and a translation vector $\boldsymbol{t}(t) \in \mathbb{R}^3$, as

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{R} & \boldsymbol{t} \\ 0 & 1 \end{pmatrix}, \qquad (2)$$

and accordingly, the points in the sensor coordinate frame $\mathcal{P}_{(q,i)}^S$ can be defined as

$$\mathcal{P}_{(q,i)}^S = {^L}\boldsymbol{T}_{S,q}(t)^{-1}\mathcal{P}_{(q,i)}^L, \qquad (3)$$

while the transformation $^L\boldsymbol{T}_{S,q}(t)$ can be extracted from

$$^L\boldsymbol{T}_{S,q}(t) = {^L}\boldsymbol{T}_{J_{pitch},q}(t)^{J_{pitch}}\boldsymbol{T}_{J_{roll},q}(t)^{J_{roll}}\boldsymbol{T}_{S,q}(t). \quad (4)$$

The geometric properties of the point cloud $\mathcal{P}_{(q,i)}^S$ in the sensor coordinate frame (i.e., roughness and normal vector) remain unchanged after performing the rigid transformations. Before the feature extraction, we use linear interpolation between the predicted pose and the previously estimated pose to remove the distortion of the point cloud similar to the approach from [22]. We extract the roughness value as a feature representation in this work. In this context, roughness value defines a term to evaluate the smoothness of the local surface. The feature extraction process is similar to the method used in [21]. However, instead of extracting the one-dimensional roughness of the points along the LiDAR's beam-lines, we extract a two-dimensional roughness value from the organized point cloud. We extract the two-dimensional roughness value by taking all neighboring points within a predefined spherical area and calculate the roughness
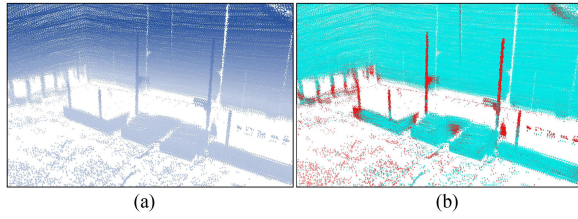
Fig. 6.   Illustration of the feature points in the LiDAR point cloud data. For better visualization, we show the accumulated CSPC data. **a)** Shows the raw point cloud from the LiDAR measurements color-coded along the z axis. **b)** Shows the calculated roughness of the points color-coded with cyan for low and red for high roughness values.

value in both vertical and horizontal directions. First, we find the set of adjacent points using the nearest neighbor search along the KD-tree. To improve the system's performance, as the size of each slice is smaller compared to the full scan mode, we filter correspondences to maintain small patches using ties broken by distance. This allows us to reduce the estimation time of the feature points for each slice. We define the two-dimensional roughness value $\mathcal{Z}$ for the point i by calculating the differences of the neighboring point set $\mathcal{N}^S_{(q,j),n}$ and normalize the term according to the distance as

$$\mathcal{Z}_i = 1 - \left\| \frac{1}{n \mathcal{P}^S_{(q,i)}} \sum_{j=1}^{n} \mathcal{N}^S_{(q,j)} \right\|, \qquad (5)$$

where n is the number of the considered nearest neighbors of point i. The feature point set $\tilde{\mathcal{D}}_{(q,i)}$ is stored with the corresponding calculated roughness value. Recalling that each slice is a small portion of one full rotation of the LiDAR, the extracted feature points, which are sorted based on their roughness value, are well distributed within the observed point cloud for the entire FoV. In addition, we use the local entropy to assess the saliency of the feature points [35], and we remove those that have a higher local entropy value. This procedure helps us to avoid selecting points that have a high roughness score while not being reliable feature points. This can be inferred from the fact that their nearby points are not well aligned. For a simple example, consider the points distributed along with the objects on a crowded table. These points might have a high roughness score, but they are not as reliable as the corner points on the wall's edge for the localization. Fig. 6 illustrates the result of the proposed method for feature extraction.

C. Low-Latency Localization and Mapping

Pose Estimation. The localization framework proposed in this letter uses the predicted pose as an initial condition. It matches the feature points of the sliced scan into the map and extracts the corresponding transformation. This procedure is concurrently performed on each of the combined sliced point sets $\mathcal{D}^S_q$. The output of this process are synchronous individual pose estimates. The combined feature point set $\mathcal{D}^S_q$ can be determined by

$$\mathcal{D}^S_q = {}^M\boldsymbol{T}_{S,q-1}^{-1} \mathcal{D}^M_{q-1} + \tilde{\mathcal{D}}_q. \qquad (6)$$

The estimated poses are later used for updating the state estimator in the EKF. Considering the point sets $\mathcal{D}^S_q$ and $\mathcal{M}^M_q$ in correspondence $\mathcal{D}^S_q \leftrightarrow \mathcal{M}^M_q$, and that they are related via a rigid body transform, in each thread we seek to estimate $\boldsymbol{R}$ and $\boldsymbol{t}$ such that

$$\mathcal{M}^M_q = \boldsymbol{R}\mathcal{D}^S_q + \boldsymbol{t} = {}^M\boldsymbol{T}_{S,q}(t)\mathcal{D}^S_q, \qquad (7)$$

where the problem is well-studied with various closed-form solutions in the literature. To solve for the pose within the map, we need to establish a geometric relationship between $\mathcal{D}^M_q$ and $\mathcal{M}^M_q$. Using the predicted transformation ${}^M\hat{\boldsymbol{T}}_{S,q}(t + \Delta q)$, which we extracted from the previous observations of the motion, we first transform the feature points in $\mathcal{D}^S_q$ into the map coordinate frame using

$$\hat{\mathcal{D}}^M_q = {}^M\hat{\boldsymbol{T}}_{S,q}(t + \Delta q)\mathcal{D}^S_q. \qquad (8)$$

Both of the point sets now are in the same coordinate frame, but they are not aligned perfectly as we used only the predicted transformation. However, we can find the nearest neighbor points for each feature point i in $\hat{\mathcal{D}}^M_q$ within the map point set. It is worth to mention that some of the points do not have any correspondence on the map. This means the closest point in the map is far from the selected feature point in $\hat{\mathcal{D}}^M_q$; therefore, we filter them using a simple threshold value based on Euclidean distance. Now we can derive a geometric relationship between all of the corresponding selected points in $\hat{\mathcal{D}}^M_q$, and the nearest points in the map as a non-linear function

$$f_{q,i}(\hat{\mathcal{D}}^M_{q,i}, {}^M\boldsymbol{T}_{S,q}) = \boldsymbol{d}_i, \quad i \in \hat{\mathcal{D}}^M_q \in \mathcal{M}^M_q, \qquad (9)$$

where stacking Eq. (9) for each feature point in $\mathcal{D}^S_q$, we can obtain a non-linear function f_q and rewrite it as

$$f_q(\mathcal{D}^S_q, {}^M\hat{\boldsymbol{T}}_{S,q}^{-1M}\boldsymbol{T}_{S,q}) = \boldsymbol{d}, \quad q \in Z^+, \qquad (10)$$

where $\boldsymbol{d}$ contains the corresponding Euclidean distances for each point, and each row of f is associated with one feature point. We use the Levenberg-Marquardt trust-region algorithm as a non-linear least squares solver for this problem [36]. In our optimization problem, a groups of scalars must be converged together. This means the three components of a translation vector and the four components of the Quaternion that define the sensor's pose are used as a parameter set. We use the Huber loss function to reduce the influence of outliers on the non-linear least squares solver's solution. To solve Eq. (10), we compute the Jacobian matrix $\boldsymbol{J}$ for the non-linear function f with respect to ${}^M\boldsymbol{T}_{S,q}$ where

$$\boldsymbol{J} = \frac{\partial f}{\partial {}^M\boldsymbol{T}_{S,q}}, \qquad (11)$$

and we further solve Eq. (10) with non-linear iteration toward minimizing the $\boldsymbol{d}$ to zero.

Linear Pose Predictor: We use a linear prediction based on the observed history of the estimated pose as an initial transformation for finding the correspondences within the map. We define the predicted transformation ${}^M\hat{\boldsymbol{T}}_{S,q}(t + \Delta q) = (\hat{\boldsymbol{R}}[\Theta^{-1}], \hat{\boldsymbol{t}})$
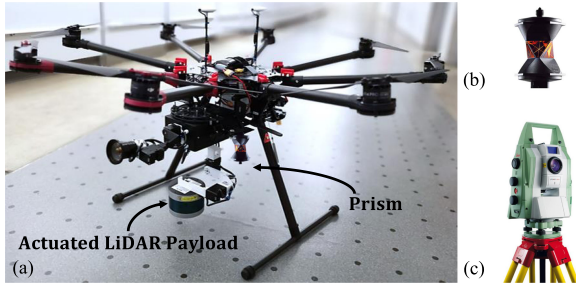
Fig. 7. (a) A UAV platform with an on-board computing unit and an active LiDAR payload. (b) The 360° reflector prism for tracking of the UAV (c) The Leica TS60 Tachymeter total station.

using the linear prediction model

$$^{M}\hat{\boldsymbol{t}}_{S,q}(t + \Delta q) = \sum_{i=0}^{b} \boldsymbol{A}_i {}^{M}\boldsymbol{t}_{S,q}(t - i\Delta q), \quad (12)$$

$$^{M}\hat{\boldsymbol{R}}_{S,q}(t + \Delta q) = {}^{M}\boldsymbol{R}_{S,q} \left[\sum_{i=0}^{b} \boldsymbol{A}_i \Theta^{-1}(t - i\Delta q) \right], \quad (13)$$

where b indicates the number of previous observations, $\boldsymbol{A}_i = \lambda_{i,b}\boldsymbol{I}_3$ is the predictor coefficients, and Θ is the element-wise Euler angle vector representation that is used to calculate the rotation matrix. Because the previously observed poses are equally-spaced values in time, a polynomial interpolation can be defined as a linear combination of the given observation. The predictor coefficient $\lambda_{i,b}$ is a scalar multiplication for the identity matrix $\boldsymbol{I}_3$. In our linear model these elements defined by the ith component of the bth row of Pascal's triangular matrix of the binomial coefficients [37].

VI. EXPERIMENTAL EVALUATION

To assess the performance of the proposed framework, we tested our approach on a recorded data set from a large sports hall. To evaluate the pose estimation accuracy, an octocopter equipped with the sensor payload is tracked with a Leica TS60 Tachymeter total station. The total station in tracking mode has a stated measurement accuracy of 0.2 cm at a rate of 10 Hz. To enable tracking with the total station, a 360° reflector prism is mounted on the octocopter as shown in Fig. 7. The static transformation is applied to have the total station's measurements in the same coordinate system as the UAV. The octocopter is flown multiple times with different trajectories in a sports hall. Flight duration in total is about 36 minutes, and the velocity varies from 0.05 to 6.0 km/h. The data were recorded applying Lissajous, raster, and static modes. In the static mode, the LiDAR was held in a horizontal state. In the other two modes, the LiDAR was actuated using the defined patterns. Five data acquisition options are used for the LiDAR point cloud; the default mode (Full) is a 360° scan at 20 Hz, and the other modes are using scan slicing with a segment size of 30, 45, 90, and 180 degrees, respectively.

An example of 6D trajectory tracking using LoLa-SLAM and the total station is illustrated in Fig. 8. The results show that
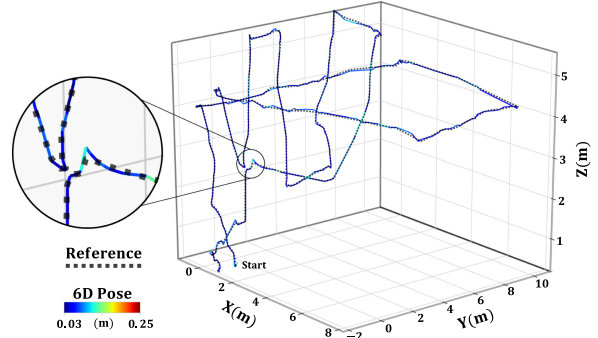
Fig. 8. Absolute trajectory error of the UAV in a sports hall in comparison to the GT pose trajectory from the total station.
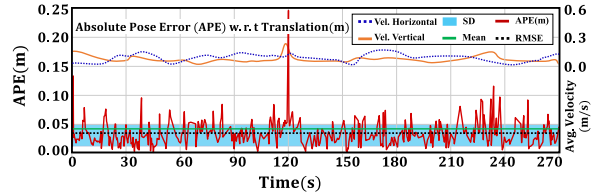
Fig. 9. Distribution of the 6D pose error with respect to the GT data of the total station. The root mean square error (RMSE) for this trajectory is equal to 3.6 cm.

TABLE I
RESULTS OF MULTIPLE EXPERIMENTS USING THE RECORDED DATA SET IN A
SPORTS HALL. ALL THE EXPERIMENTS ARE CONDUCTED USING THE ONBOARD
CORE.I7 INTEL PROCESSOR - 4 CORES AND WITH 16 GB OF RAM

	Experiment ($\theta°$ + Pattern)	Total Dist. (m)	CPU Usage Avg.	RMSE - R (°)	RMSE - T (m)	Max Error (m)	SD (m)	Update Rate (Hz)	Latency (ms)
LoLa-SLAM (Proposed)	30°+Lissajous	184	-	-	**Failed**	-	-	-	-
	30°+Raster	175	-	-	**Failed**	-	-	-	-
	30°+Static	193	92%	7.4	**0.079**	0.48	0.041	240	16
	45°+Lissajous	184	<u>85%</u>	4.3	<u>**0.039**</u>	0.27	0.028	<u>160</u>	<u>19</u>
	45°+Raster	175	83%	4.4	**0.057**	0.41	0.033	160	21
	45°+Static	193	77%	5.8	**0.070**	0.42	0.034	160	18
	90°+Lissajous	184	84%	4.4	**0.036**	0.27	0.033	80	39
	90°+Raster	175	83%	4.4	**0.055**	0.43	0.032	80	43
	90°+Static	193	75%	5.8	**0.067**	0.46	0.034	80	38
	180°+Lissajous	184	55%	4.1	**0.035**	0.25	0.027	40	78
	180°+Raster	175	51%	4.9	**0.052**	0.39	0.029	40	85
	180°+Static	193	52%	3.7	**0.066**	0.38	0.031	40	77
LOAM	Full+Lissajous	184	<u>53%</u>	4.6	<u>**0.034**</u>	0.26	0.024	<u>20</u>	<u>93</u>
	Full+Raster	175	56%	4.8	**0.050**	0.44	0.031	20	97
	Full+Static	193	49%	3.9	**0.065**	0.47	0.033	20	89
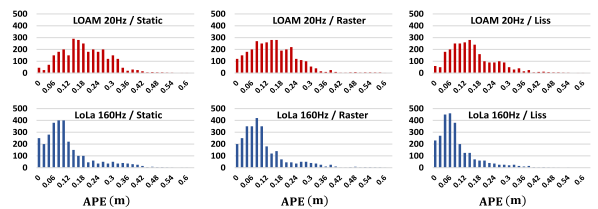
Fig. 10. Histogram of the absolute pose error (APE) distribution when performing active-controlled hovering (PD controller) with 20 Hz and 160 Hz pose feedback.
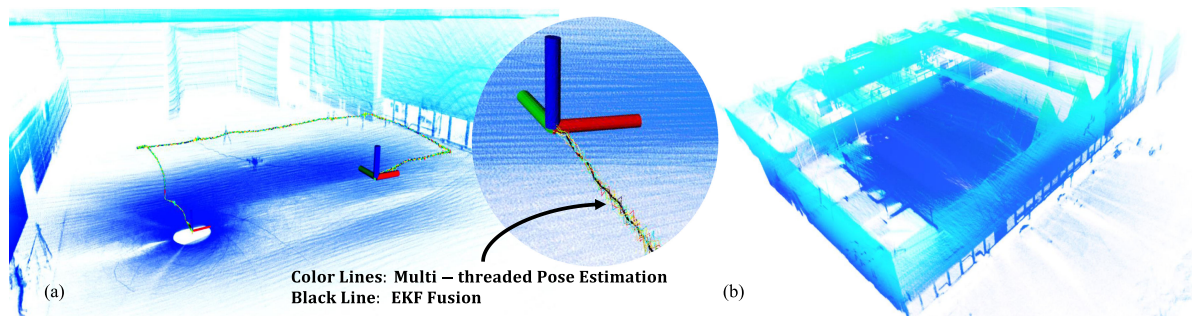
Fig. 11.   Reconstructed 3D map of a sports hall using the LoLa-SLAM framework. *a*) Concurrent pose estimations from the multi-threaded localization and the fused 6D pose of the UAV. *b*) The dense point cloud map from a different perspective.

there is less deviation in the horizontal direction compared to the vertical direction. We can see that the larger errors mostly appear when the drone moves fast in the vertical direction. This issue could have two origins; First, for the linear predictor, as the map matching uses this prediction as an initial input in the localization step, a poor prediction could cause such a wrong matching. Second, imperfect observations caused by the LiDAR actuation: due to the insufficient points in the vertical face compared to the horizontal. To investigate more on this problem, we further compared the proposed method with the static LiDAR; this means even fewer points are observed in the vertical faces, and as expected, the error in the vertical direction increased considerably. In addition to this, as shown in Fig. 9, the RMSE of a selected trajectory is only 3.6 cm; however, due to the fast motion, the maximum pose error increased up to 26.8 cm at some points.

Furthermore, the latency and accuracy of the pose estimation are investigated. In Table I, we provide a comparison to the well known state-of-the-art SLAM framework LOAM [21]. By analyzing the results, we can see a trade-off between the chosen slice size, pose estimation accuracy, and the update rate. Choosing a smaller slice size results in a higher update rate; however, this means fewer features in each slice are used for localization. Consequently, the accuracy of the pose estimation starts deviating. In an extreme scenario, with using the slice size $\theta = 30°$, the system cannot track the pose, and therefore the localization starts drifting. If the drift is larger than 0.5 m, we consider the localization to fail.

Considering a slice size $\theta = 45°$ is a satisfactory candidate, we achieve a latency of less than 20 ms on average, which is acceptable for real-time drone navigation. In comparison, using the full scan mode, we observe a latency of 93 ms. Besides, we investigate the use of LoLa-SLAM in a closed-loop system while the drone was hovering using a PD controller. The histogram of the APE distribution is shown in Fig. 10. The figure shows that, while the accuracy of the pose estimation is lower using LoLa-SLAM compared to LOAM, the drone deviation is smaller due to the shorter latency and higher temporal update rate. While LoLa-SLAM provides a higher update rate and lower latency, it is computationally more expensive than the other approaches. For instance, on average, the proposed framework needs around 30% more processing power compared to LOAM. However, thanks
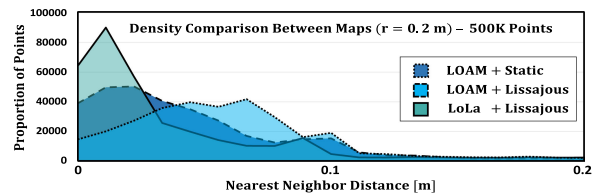
Fig. 12.   Density comparison between the 3D maps. The distribution of the point spacing (nearest-neighbor distance) for randomly selected 500 K points within each map.

to the concurrent multi-threaded map matching, the system still runs in real-time.

The accuracy and overall validity of the proposed system are additionally assessed using the created dense 3D map. To evaluate the map quality, Fig. 11 shows the resulting point cloud from different perspectives, which allows for better interpretation of the scene. The tripod of the total station and pillars in the sports hall can be identified in the globally aligned and voxelized 3D map. It is worth stating that this level of 3D mapping is not obtainable without a precise localization. To evaluate the superior quality of the created map compared to LOAM, we further illustrate the density comparison between the point clouds similar to the method described in [25]. The point densities show the uniformity of the generated map. The distribution of the point spacing (nearest-neighbor distance) for randomly samples points is visualized in Fig. 12. The densities are represented as distributions of point spacing, as determined by each point's nearest neighbor. We use randomly selected 500 K points within each map and determined the density in a sphere with 0.2 m area around each point. This comparison shows that LoLa-SLAM has the highest proportion of points in the distance of less than 3.0 cm.

VII. CONCLUSION

In this work, we presented a customized system to localize a UAV in an indoor GPS-denied environment. We approached this challenge by employing an actuated LiDAR platform and utilizing the 2D Lissajous spinning pattern. We further developed a novel SLAM framework based on the continuous sliced point cloud to reduce the ego-motion estimation latency, named

LoLa-SLAM. We showed that the proposed SLAM system's accuracy and reliability remain adequate, and this is while the temporal update rate increased by at least 2X for the pose estimation. Besides, we demonstrated the proposed framework could simultaneously generate a high-quality 3D map. We demonstrate both the low-latency pose estimation and the overall accuracy of LoLa-SLAM by multiple experiments in an indoor environment. We showed that the update rate increases by choosing a smaller slice size; however, the proposed system can fail in the feature matching process if the slice size is too small. In addition, we examined the computational complexity which increases by 30% compared to the common methods.

In future work, to further improve the system, the odometry predictor and the EKF fusion can be integrated with redundant attitude and heading reference systems (AHRS) [24], [38]. Furthermore, adding an adaptive slice size controller based on the UAV's velocity can also be investigated to improve the system performance. Besides, environment aware attentive control of the LiDAR can be employed to improve the point cloud observation, map generation, and reliable pose estimation in complex indoor environments.

REFERENCES

[1] J. Miranda, S. Larnier, A. Herbulot, and M. Devy, "Uav-based inspection of airplane exterior screws with computer vision," in *Proc. 14th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, Feb. 2019.

[2] F. Valenti, D. Giaquinto, L. Musto, A. Zinelli, M. Bertozzi, and A. Broggi, "Enabling computer vision-based autonomous navigation for unmanned aerial vehicles in cluttered gps-denied environments," in *Proc. Int. Conf. Intell. Transp. Syst.*, 2018, pp. 3886–3891.

[3] J. Siva and C. Poellabauer, "Robot and drone localization in gps-denied areas," in *Mission-Oriented Sensor Networks and Systems: Art and Science*. Berlin, Germany: Springer, 2019, pp. 597–631.

[4] Y. D. Yasuda, L. E. G. Martins, and F. A. Cappabianco, "Autonomous visual navigation for mobile robots: A systematic literature review," *ACM Comput. Surv.*, vol. 53, no. 1, pp. 1–34, 2020.

[5] X. Zhou, Z. Yi, Y. Liu, K. Huang, and H. Huang, "Survey on path and view planning for uavs," *Virtual Reality Intell. Hardware*, vol. 2, no. 1, pp. 56–69, 2020.

[6] S. Jung, S. Hwang, H. Shin, and D. H. Shim, "Perception, guidance, and navigation for indoor autonomous drone racing using deep learning," *IEEE Robot. Automat. Lett.*, vol. 3, no. 3, pp. 2539–2544, Jul. 2018.

[7] R. S. Dimitrova, M. Gehrig, D. Brescianini, and D. Scaramuzza, "Towards low-latency high-bandwidth control of quadrotors using event cameras," *IEEE Int. Conf. Robot. Autom. (ICRA)*, 2020, pp. 4294–4300.

[8] R. Pérez-Alcocer and J. Moreno-Valenzuela, "A novel lyapunov-based trajectory tracking controller for a quadrotor: Experimental analysis by using two motion tasks," *Mechatronics*, vol. 61, pp. 58–68, 2019.

[9] R. Mur-Artal and J. D. Tardós, "Visual-inertial monocular SLAM with map reuse," *IEEE Robot. Automat. Lett.*, vol. 2, no. 2, pp. 796–803, Apr. 2017.

[10] D. Van Opdenbosch, M. Oelsch, A. Garcea, T. Aykut, and E. Steinbach, "Selection and compression of local binary features for remote visual SLAM," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 7270–7277.

[11] S. Lynen *et al.*, "Large-scale, real-time visual-inertial localization revisited," *Int. J. Robot. Res.*, vol. 39, no. 9, pp. 1061–1084, 2020.

[12] M. W. Achtelik, S. Lynen, S. Weiss, L. Kneip, M. Chli, and R. Siegwart, "Visual-inertial SLAM for a small helicopter in large outdoor environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2012, pp. 2651–2652.

[13] D. Zhang, Z. Gong, Y. Chen, J. Zelek, and J. Li, "SLAM-based multi-sensor backpack lidar systems in gnss-denied environments," in *Proc. Geosci. Remote Sens. Symp.*, 2019, pp. 8984–8987.

[14] M. Velas, M. Spanel, T. Sleziak, J. Habrovec, and A. Herout, "Indoor and outdoor backpack mapping with calibrated pair of velodyne lidars," *Sensors*, vol. 19, no. 18, 2019, Art. no. 3944.

[15] J. Lin and F. Zhang, "Loam_livox: A fast, robust, high-precision lidar odometry and mapping package for lidars of small fov," *IEEE Int. Conf. Robot. Autom. (ICRA)*, 2020, pp. 3126–3131.

[16] J. Lin and F. Zhang, "A fast, complete, point cloud based loop closure for lidar odometry and mapping," 2019, *arXiv:1909.11811*.

[17] H. Qin *et al.*, "A stereo and rotating laser framework for UAV navigation in GPS denied environment," in *Proc. IECON 42nd Annu. Conf. IEEE Ind. Electron. Soc.*, 2016, pp. 6061–6066.

[18] W. Zhen and S. Scherer, "A unified 3D mapping framework using a 3D or 2D lidar," *Int. Symp. Experimental Robot.*, 2018, pp. 702–711.

[19] H. Qin *et al.*, "Autonomous exploration and mapping system using heterogeneous uavs and ugvs in gps-denied environments," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1339–1350, Feb. 2019.

[20] R. Voges, C. S. Wieghardt, and B. Wagner, "Timestamp offset determination between an actuated laser scanner and its corresponding motor," *ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 4, pp. 99–106, 2017.

[21] J. Zhang and S. Singh, "Loam: Lidar odometry and mapping in real-time," *Robot.: Sci. Syst.*, vol. 2, no. 9, 2014.

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

[23] X. Ji, L. Zuo, C. Zhang, and Y. Liu, "Lloam: Lidar odometry and mapping with loop-closure detection based correction," in *Proc. IEEE Int. Conf. Mechatronics Automat.*, 2019, pp. 2475–2480.

[24] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and D. Rus, "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping," 2020, *arXiv:2007.00258*.

[25] L. Kaul, R. Zlot, and M. Bosse, "Continuous-time three-dimensional mapping for micro aerial vehicles with a passively actuated rotating laser scanner," *J. Field Robot.*, vol. 33, no. 1, pp. 103–132, 2016.

[26] S. Anderson and T. D. Barfoot, "Towards relative continuous-time SLAM," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2013, pp. 1033–1040.

[27] M. Bosse, R. Zlot, and P. Flick, "Zebedee: Design of a spring-mounted 3D range sensor with application to mobile mapping," *IEEE Trans. Robot.*, vol. 28, no. 5, pp. 1104–1119, Oct. 2012.

[28] J. W. Anderson and G. M. Clayton, "Lissajous-like scan pattern for a gimballed lidar," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics*, 2014, pp. 1171–1176.

[29] M. Benson, J. Nikolaidis, and G. M. Clayton, "Lissajous-like scan pattern for a nodding multi-beam lidar," in *Proc. ASME Dyn. Syst. Control Conf.*, 2018, pp. 1171–1176.

[30] D. Droeschel, M. Nieuwenhuisen, M. Beul, D. Holz, J. Stückler, and S. Behnke, "Multilayered mapping and navigation for autonomous micro aerial vehicles," *J. Field Robot.*, vol. 33, no. 4, pp. 451–475, 2016.

[31] W. Zhen, S. Zeng, and S. Soberer, "Robust localization and localizability estimation with a rotating laser scanner," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2017, pp. 6240–6245.

[32] H. Nakagomi *et al.*, "3D scan matching for mobile robot localization over rough terrain," *Elect. Eng. Jpn.*, vol. 209, no. 3/4, pp. 14–25, 2019.

[33] M. T. Benson, H. Sathishchandra, G. M. Clayton, and S. B. Andersson, "Compressive sensing-based reconstruction of lissajous-like nodding lidar data," in *Proc. Dyn. Syst. Control Conf.*, 2019, Art. no. V003T21A010.

[34] A. Al-Nuaimi, W. Lopes, P. Zeller, A. Garcea, C. Lopes, and E. Steinbach, "Analyzing lidar scan skewing and its impact on scan matching," in *Proc. IEEE Int. Conf. Indoor Positioning Indoor Navigation*, 2016, pp. 1–8.

[35] G. Li, Y. Geng, and W. Zhang, "Autonomous planetary rover navigation via active SLAM," *Aircraft Eng. Aerospace Tech.*, vol. 91, no. 1, pp. 60–68, 2018.

[36] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[37] C. K. Williams, "Prediction with gaussian processes: From linear regression to linear prediction and beyond," in *Learning in Graphical Models*. Berlin, Germany: Springer, 1998, pp. 599–621.

[38] M. Karimi, E. Babaians, M. Oelsch, T. Aykut, and E. Steinbach, "Skewed-redundant hall-effect magnetic sensor fusion for perturbation-free indoor heading estimation," in *Proc. 4th IEEE Int. Conf. Robot. Comput.*, 2020, pp. 367–374.

Chapter 6

Conclusion and Future Work

Through the integration of range, magnetic, and inertial sensing, this thesis considered how to estimate real-time and low-latency 6D ego-motion while constructing a point-cloud map of the explored environment. The focus of this work is divided into two main sections; robust attitude and heading estimation in an indoor environment, and low-latency 6D ego-motion estimation and mapping techniques. In the first chapter of this thesis, we introduced the problem. In Chapter 2, we provide the fundamentals necessary to understand the thesis's core ideas. In addition, we reviewed the literature related to magnetic, inertial, and laser based methods that have been utilized in the field of 6D SLAM.

In the context of Chapter 3 and 4, we presented a fusion and filtering approach for robust attitude and heading estimation in an indoor environment. This approach combined skewed magnetometers and inertial sensors. We demonstrated heading estimation by using the Earth's magnetic field and the different sources of perturbation for heading angle estimation in the presence of external magnetic anomalies. Using a three-axis MEMS sensor, we investigated the Hall-effect magnetic sensor's principle, as well as analyzed the external magnetic field's source. We designed a tetrahedron platform based on the skewed redundant configuration, using triple 9-axes sensors to detect and mitigate over-saturated channels in the observation signals. A correlation-based filter model for multiple magnetometer platforms has been presented. For the skewed redundant MIMU fusion, RNN-based fusion models were presented. We investigated multiple fusion models using LSTM-based neural networks. Additionally, we investigated how the models performed in different indoor environments with different scenarios, and validated the algorithm by using the data we had recorded internally. Moreover, experiments using a gimbal platform have been conducted in a number of scenarios to demonstrate robust attitude and heading estimation using the skewed redundant platform to verify the generality of the proposed fusion method. The experimental results indicate that although the proposed system is reliable for estimating heading over long-term periods, it is extremely sensitive to the size of the input measurement history. As a result, for short-term heading estimation, conventional methods performed similar or better than the developed model. However, due to the observed drifts in long-term use, the proposed system provides a more reliable estimation with less deviation in a saturated indoor environment. The proposed framework achieved a Root Mean Square

Error in attitude and heading estimation of less than 2.5 degrees for long-term use. Finally, the result of the attitude and heading estimation is further utilized in the next Chapter where we tackle the 6D low-latency ego-motion estimation.

In Chapter 5, we present a 6D SLAM framework with low latency ego-motion estimation in an indoor environment that has GPS denial. We approached the 6D pose estimation challenge by employing an actuated LiDAR platform and utilizing the 2D Lissajous spinning pattern. We further developed a novel framework based on the continuous sliced point cloud to reduce the ego-motion estimation latency. The experimental results showed that the proposed SLAM system's accuracy and reliability remained excellent, even when the temporal update rate increased by 4x for the estimation of ego-motion. In addition, we showed that the proposed framework could be used to construct a 3D point cloud map from the captured data of the explored area. In multiple experiments in an indoor environment, we demonstrate both the low-latency ego-motion estimation as well as the overall accuracy of the LiDAR-MIMU SLAM algorithm. The proposed framework enables the 6D pose estimation in less than 20ms latency. Generally, we found that the update rate increases when choosing a smaller slice size, however, the proposed system can fail in the feature matching process when the slice size is too small. A further consideration was the computational complexity, which increased by 30% when compared with the conventional methods.

As the result of this work and state-of-the-art, skewed redundant observation of magnetic and intertial sensors enables significant improvement in final state estimation. The effect of the number of sensing axes on the final state estimation has been established. Specially, with the power of DNN, complex fusion models can be applied regardless of the degree of measurement redundancy. To this end, in future work, to enhance the accuracy of the fusion approach for attitude and heading reference estimation, continuous-redundant MIMU systems can be investigated. The continuous-redundant MIMU idea emerged when we investigated the actuated LiDAR platform for solving the 6D pose estimation problem. In this way, a single 9-DoF MIMU can monitor a larger FOV (or measurement axis in this case), where the DNN can be used to fuse the observation to the final estimate of attitude and heading reference. As of the date of presenting this work, there are no redundant-axes, or continuous-axis MEMS-MIMU systems available in the market. Sensors with redundant-axes or continuous-axes can utilize a MEMS-based or beam steering technique to observe acceleration, magnetic field, or angular velocity from multiple axes in a small footprint. Such a system can be a perfect fit for the future research direction in this field.

Additionally, the event-based LiDAR sensors can compress the information of moving objects in a cost effective way, which in turn, enables energy-efficient and real-time processing in various applications such as object detection, motion recognition, and SLAM. The possibility of achieving superior performance in terms of accuracy and latency by using event-based measurement techniques can be explored for future work. In this regard, as to this date where this work is presented, there is no event-based LiDAR system available in the market,

however, enhancing LiDAR point clouds with event-based cameras would be the perfect fit for future research directions. Event-based cameras react to the changes in light intensity and output dense event streams consisting of triggered pixels. Unfortunately, depth information is not available for event cameras. In order to address these problems, enhancing LiDAR and event-based cameras would lead to a significant reduction in ego-motion estimation latency as well as an increase in the overall performance of such a multi-sensor system.

Bibliography

Publications by the author

Journal publications

[1] M. Karimi, E. Babaians, M. Oelsch, and E. Steinbach, "Deep fusion of a skewed redundant magnetic and inertial sensor for heading state estimation in a saturated indoor environment," *International Journal of Semantic Computing*, vol. 15, no. 03, pp. 313–335, 2021. DOI: 10.1142/S1793351X21400079.

[2] M. Karimi, M. Oelsch, O. Stengel, E. Babaians, and E. Steinbach, "Lola-slam: Low-latency lidar slam using continuous scan slicing," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2248–2255, 2021. DOI: 10.1109/LRA.2021.3060721.

[3] M. Oelsch, M. Karimi, and E. Steinbach, "R-loam: Improving lidar odometry and mapping with point-to-mesh features of a known 3d reference object," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2068–2075, 2021. DOI: 10.1109/LRA.2021. 3060413.

[4] ——, "Ro-loam: 3d reference object-based trajectory and map optimization in lidar odometry and mapping," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 6806–6813, 2022. DOI: 10.1109/LRA.2022.3177846.

[5] T. Aykut, M. Karimi, C. Burgmair, A. Finkenzeller, C. Bachhuber, and E. Steinbach, "Delay compensation for a telepresence system with 3d 360 degree vision based on deep head motion prediction and dynamic fov adaptation," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4343–4350, 2018. DOI: 10.1109/LRA.2018.2864359.

Conference publications

[6] M. Karimi, E. Babaians, M. Oelsch, T. Aykut, and E. Steinbach, "Skewed-redundant hall-effect magnetic sensor fusion for perturbation-free indoor heading estimation," in *2020 Fourth IEEE International Conference on Robotic Computing (IRC)*, IEEE, 2020, pp. 367–374. DOI: 10.1109/IRC.2020.00064.

[7] M. Karimi, T. Aykut, and E. Steinbach, "Mavi: A research platform for telepresence and teleoperation," *arXiv preprint arXiv:1805.09447*, 2018. DOI: 10.48550/arXiv.1805. 09447.

[8]   M. Oelsch, M. Karimi, and E. Steinbach, "Init-loam: Lidar-based localization and mapping with a static self-generated initial map," in *20th International Conference on Advanced Robotics (ICAR)*, IEEE, 2021, pp. 865–872. DOI: 10.1109 / ICAR53236.2021. 9659358.

[9]   M. Tappe, D. Dose, M. Oelsch, *et al.*, "Uas based autonomous visual inspection of airplane surface defects," in *NDE 4.0, Predictive Maintenance, and Communication and Energy Systems in a Globally Networked World*, SPIE, vol. 12049, 2022, pp. 8–21. DOI: 10.1117/12.2612579.

General publications

[10]   T. Bailey, J. Nieto, J. Guivant, M. Stevens, and E. Nebot, "Consistency of the ekf-slam algorithm," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2006, pp. 3562–3568.

[11]   J. Kelly and G. S. Sukhatme, "Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration," *The International Journal of Robotics Research*, vol. 30, no. 1, pp. 56–79, 2011.

[12]   T. Tomic, K. Schmid, P. Lutz, *et al.*, "Toward a Fully Autonomous UAV: Research Platform for Indoor and Outdoor Urban Search and Rescue," *IEEE Robotics Automation Magazine*, vol. 19, no. 3, pp. 46–56, Sep. 2012, ISSN: 1558-223X. DOI: 10.1109 / MRA. 2012.2206473.

[13]   V. Kangunde, R. S. Jamisola, and E. K. Theophilus, "A review on drones controlled in real-time," *International journal of dynamics and control*, vol. 9, no. 4, pp. 1832–1846, 2021.

[14]   S. Winkvist, E. Rushforth, and K. Young, "Towards an autonomous indoor aerial inspection vehicle," *Industrial Robot: An International Journal*, vol. 40, no. 3, pp. 196–207, Jan. 2013, ISSN: 0143-991X. DOI: 10.1108/01439911311309870. [Online]. Available: https://doi.org/10.1108/01439911311309870.

[15]   R. F. Brena, J. P. García-Vázquez, C. E. Galván-Tejada, D. Muñoz-Rodriguez, C. Vargas-Rosales, and J. Fangmeyer, "Evolution of indoor positioning technologies: A survey," *Journal of Sensors*, vol. 2017, Mar. 2017. DOI: 10.1155/2017/2630413.

[16]   A. Yassin, Y. Nasser, M. Awad, *et al.*, "Recent Advances in Indoor Localization: A Survey on Theoretical Approaches and Applications," *IEEE Communications Surveys Tutorials*, vol. 19, no. 2, pp. 1327–1346, 2017, ISSN: 2373-745X. DOI: 10.1109/COMST. 2016.2632427.

[17]   A. Benini, A. Mancini, and S. Longhi, "An IMU/UWB/Vision-based Extended Kalman Filter for Mini-UAV Localization in Indoor Environment using 802.15.4a Wireless Sensor Network," *Journal of Intelligent & Robotic Systems*, vol. 70, no. 1, pp. 461–476, 2013, ISSN: 1573-0409. DOI: 10.1007/s10846-012-9742-1. [Online]. Available: https://doi.org/10.1007/s10846-012-9742-1.

[18] J. Tiemann, F. Schweikowski, and C. Wietfeld, "Design of an UWB indoor-positioning system for UAV navigation in GNSS-denied environments," in *2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Oct. 2015, pp. 1–7. DOI: 10.1109/IPIN.2015.7346960.

[19] A. Masiero, F. Fissore, and A. Vettore, "A Low Cost UWB Based Solution for Direct Georeferencing UAV Photogrammetry," *Remote Sensing*, vol. 9, no. 5, p. 414, 2017, ISSN: 2072-4292. DOI: 10.3390/rs9050414. [Online]. Available: http://dx.doi.org/10.3390/rs9050414.

[20] M. Zhou, J. Lin, S. Liang, W. Du, and L. Cheng, "A UAV patrol system based on Bluetooth localization," in *2017 2nd Asia-Pacific Conference on Intelligent Robot Systems (ACIRS)*, Jun. 2017, pp. 205–209. DOI: 10.1109/ACIRS.2017.7986094.

[21] P. R. Soria, A. F. Palomino, B. C. Arrue, and A. Ollero, "Bluetooth network for micro-uavs for communication network and embedded range only localization," in *2017 International Conference on Unmanned Aircraft Systems (ICUAS)*, Jun. 2017, pp. 747–752. DOI: 10.1109/ICUAS.2017.7991464.

[22] A. Ollero and L. Merino, "Control and perception techniques for aerial robotics," *Annual Reviews in Control*, vol. 28, no. 2, pp. 167–178, 2004, ISSN: 1367-5788. DOI: https://doi.org/10.1016/j.arcontrol.2004.05.003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1367578804000367.

[23] D. Dey, K. S. Shankar, S. Zeng, *et al.*, "Vision and Learning for Deliberative Monocular Cluttered Flight," in *Field and Service Robotics: Results of the 10th International Conference*, D. S. Wettergreen and T. D. Barfoot, Eds. Cham: Springer International Publishing, 2016, pp. 391–409, ISBN: 978-3-319-27702-8. DOI: 10.1007/978-3-319-27702-8_26. [Online]. Available: https://doi.org/10.1007/978-3-319-27702-8_26.

[24] P. Schmuck and M. Chli, "Multi-UAV collaborative monocular SLAM," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, May 2017, pp. 3863–3870. DOI: 10.1109/ICRA.2017.7989445.

[25] R. Strydom, S. Thurrowgood, and M. V. Srinivasan, "Visual odometry: Autonomous uav navigation using optic flow and stereo," in *Proceedings of Australasian conference on robotics and automation*, 2014. [Online]. Available: https://espace.library.uq.edu.au/view/UQ:356412.

[26] I. Cvišić, J. Ćesić, I. Marković, and I. Petrović, "SOFT-SLAM: Computationally efficient stereo visual simultaneous localization and mapping for autonomous unmanned aerial vehicles," *Journal of Field Robotics*, vol. 35, no. 4, pp. 578–595, 2018. DOI: 10.1002/rob.21762. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/rob.21762. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.21762.

[27] M. C. P. Santos, L. V. Santana, A. S. Brandão, and M. Sarcinelli-Filho, "UAV obstacle avoidance using RGB-D system," in *2015 International Conference on Unmanned Aircraft Systems (ICUAS)*, Jun. 2015, pp. 312–319. DOI: 10.1109/ICUAS.2015.7152305.

[28] W. G. Aguilar, G. A. Rodríguez, L. Álvarez, S. Sandoval, F. Quisaguano, and A. Limaico, "Visual SLAM with a RGB-D Camera on a Quadrotor UAV Using on-Board Processing," in *Advances in Computational Intelligence*, I. Rojas, G. Joya, and A. Catala, Eds., Cham: Springer International Publishing, 2017, pp. 596–606, ISBN: 978-3-319-59147-6. DOI: 10.1007/978-3-319-59147-6_51.

[29] Y. Lu, Z. Xue, G.-S. Xia, and L. Zhang, "A survey on vision-based UAV navigation," *Geo-spatial Information Science*, vol. 21, no. 1, pp. 21–32, 2018. DOI: 10.1080/10095020.2017.1420509. eprint: https://doi.org/10.1080/10095020.2017.1420509. [Online]. Available: https://doi.org/10.1080/10095020.2017.1420509.

[30] E. Balestrieri, P. Daponte, L. De Vito, F. Picariello, and I. Tudosa, "Sensors and measurements for uav safety: An overview," *Sensors*, vol. 21, no. 24, p. 8253, 2021. DOI: 10.3390/s21248253.

[31] A. Wilson, A. Kumar, A. Jha, and L. R. Cenkeramaddi, "Embedded sensors, communication technologies, computing platforms and machine learning for uavs: A review," *IEEE Sensors Journal*, 2021. DOI: 10.1109/JSEN.2021.3139124.

[32] J. Duník, S. K. Biswas, A. G. Dempster, T. Pany, and P. Closas, "State estimation methods in navigation: Overview and application," *IEEE Aerospace and Electronic Systems Magazine*, vol. 35, no. 12, pp. 16–31, 2020.

[33] J. S. Bennett, B. E. Vyhnalek, H. Greenall, *et al.*, "Precision magnetometers for aerospace applications: A review," *Sensors*, vol. 21, no. 16, p. 5568, 2021. DOI: 10.3390/s21165568.

[34] R. Hansenw, R. L. Kleinberg, and A. Kaufman, *Principles of the Magnetic Methods in Geophysics, Volume 42 (Methods in Geochemistry and Geophysics)*. Elsevier, 2009.

[35] S. Tumanski, *Handbook of magnetic measurements*. CRC press, 2016.

[36] B. Fan, Q. Li, and T. Liu, "How magnetic disturbance influences the attitude and heading in magnetic and inertial sensor-based orientation estimation," *Sensors*, vol. 18, no. 1, p. 76, 2017. DOI: 10.3390/s18010076.

[37] J. H. Migueles, C. Cadenas-Sanchez, U. Ekelund, *et al.*, "Accelerometer data collection and processing criteria to assess physical activity and other outcomes: A systematic review and practical considerations," *Sports medicine*, vol. 47, no. 9, pp. 1821–1845, 2017. DOI: 10.1007/s40279-017-0716-0.

[38] G. G. Samatas and T. P. Pachidis, "Inertial measurement units (imus) in mobile robots over the last five years: A review," *Designs*, vol. 6, no. 1, p. 17, 2022. DOI: 10.3390/designs6010017.

[39] H. Al-Jlailaty and M. M. Mansour, "Efficient attitude estimators: A tutorial and survey," *Journal of Signal Processing Systems*, pp. 1–35, 2021. DOI: 10.1007/s11265-020-01620-4.

[40] N. El-Sheimy and A. Youssef, "Inertial sensors technologies for navigation applications: State of the art and future trends," *Satellite Navigation*, vol. 1, no. 1, pp. 1–21, 2020. DOI: 10.1186/s43020-019-0001-5.

[41] X. Zhai, Y. Ren, L. Wang, T. Zhu, Y. He, and B. Lv, "A review of redundant inertial navigation technology," in *2021 International Conference on Computer, Control and Robotics (ICCCR)*, IEEE, 2021, pp. 272–278. DOI: 10.1109/ICCCR49711.2021.9349397.

[42] D. Korchev, S. Cheng, Y. Owechko, *et al.*, "On real-time lidar data segmentation and classification," in *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV)*, The Steering Committee of The World Congress in Computer Science, 2013. [Online]. Available: https://www.semanticscholar.org/paper/On-Real-Time-LIDAR-Data-Segmentation-and-Korchev-Cheng/5b38633aa8c09fd66f2a1ab28660cb1a8ff5e8ab.

[43] S. Grzonka, G. Grisetti, and W. Burgard, "A Fully Autonomous Indoor Quadrotor," *IEEE Transactions on Robotics*, vol. 28, no. 1, pp. 90–100, Feb. 2012, ISSN: 1941-0468. DOI: 10.1109/TRO.2011.2162999.

[44] D. Droeschel, J. Stückler, and S. Behnke, "Local multi-resolution representation for 6D motion estimation and mapping with a continuously rotating 3D laser scanner," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, May 2014, pp. 5221–5226. DOI: 10.1109/ICRA.2014.6907626.

[45] W. Zhen, S. Zeng, and S. Soberer, "Robust localization and localizability estimation with a rotating laser scanner," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, May 2017, pp. 6240–6245. DOI: 10.1109/ICRA.2017.7989739.

[46] C. Green, B. Stevens, S. Biaz, and R. Chapman, "UAV SLAM and Interior Modeling with 3D LiDAR in GNSS-Denied Environments," 2019. [Online]. Available: https://www.semanticscholar.org/paper/UAV-SLAM-and-Interior-Modeling-with-3-D-LiDAR-in/961d7ead164ae361d2d0a253c5551db419ea48d9.

[47] R. Harle, "A survey of indoor inertial positioning systems for pedestrians," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1281–1293, 2013. DOI: 10.1109/SURV.2012.121912.00075.

[48] O. S. Salychev, *Applied Inertial Navigation: problems and solutions*. BMSTU press Moscow, Russia: 2004.

[49] K. Gade, "The seven ways to find heading," *The Journal of Navigation*, vol. 69, no. 5, pp. 955–970, 2016. DOI: 10.1017/S0373463316000096.

[50] M. Sturza, "Skewed axis inertial sensor geometry for optimal performance," in *Digital Avionics Systems Conference*, 1988, p. 3874. DOI: 10.2514/6.1988-3874.

[51] A. Osman, B. Wright, A. Noureldin, and N. El-Sheimy, "Multi-sensor inertial navigation systems employing skewed redundant inertial sensors," in *ION GNSS 19th International Technical Meeting of the Satellite Division*, 2006, pp. 26–29.

[52] M. Kok and T. B. Schön, "Magnetometer calibration using inertial sensors," *IEEE Sensors Journal*, vol. 16, no. 14, pp. 5679–5689, 2016.

[53] J. F. Vasconcelos, G Elkaim, C. Silvestre, P Oliveira, and B. Cardeira, "Geometric approach to strapdown magnetometer calibration in sensor frame," *IEEE Transactions on Aerospace and Electronic systems*, vol. 47, no. 2, pp. 1293–1306, 2011.

[54] M. Kok, J. D. Hol, T. B. Schön, F. Gustafsson, and H. Luinge, "Calibration of a magnetometer in combination with inertial sensors," in *2012 15th International Conference on Information Fusion*, IEEE, 2012, pp. 787–793.

[55] J. Lenz and S Edelstein, "Magnetic sensors and their applications," *IEEE Sensors journal*, vol. 6, no. 3, pp. 631–649, 2006. DOI: 10.1109/JSEN.2006.874493.

[56] M. Kok and T. B. Schön, "Maximum likelihood calibration of a magnetometer using inertial sensors," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 92–97, 2014.

[57] H. E. Soken, "A survey of calibration algorithms for small satellite magnetometers," *Measurement*, vol. 122, pp. 417–423, 2018. DOI: 10.1016/j.measurement.2017.10.017.

[58] K. Papafotis, D. Nikitas, and P. P. Sotiriadis, "Magnetic field sensors' calibration: Algorithms' overview and comparison," *Sensors*, vol. 21, no. 16, p. 5288, 2021. DOI: 10.3390/s21165288.

[59] K. Han, H. Han, Z. Wang, and F. Xu, "Extended kalman filter-based gyroscope-aided magnetometer calibration for consumer electronic devices," *IEEE Sensors Journal*, vol. 17, no. 1, pp. 63–71, 2016. DOI: 10.1109/JSEN.2016.2624821.

[60] X. Tu, *Magnetometer calibration*, US Patent 8,717,009, 2014.

[61] P. L. Worcester, D. J. Maxwell, and M. L. Trethewey, *Calibration of magnetic compasses*, US Patent 5,117,375, 1992.

[62] J. Fang, H. Sun, J. Cao, X. Zhang, and Y. Tao, "A novel calibration method of magnetic compass based on ellipsoid fitting," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 6, pp. 2053–2061, 2011.

[63] M. Zhu, W. Ouyang, and Y. Wu, "Orientation estimation by partial-state updating kalman filter and vectorial magnetic interference detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 3, pp. 1815–1826, 2021. DOI: 10.1109/TAES.2021.3050657.

[64] G. Ouyang and K. Abed-Meraim, "A survey of magnetic-field-based indoor localization," *Electronics*, vol. 11, no. 6, p. 864, 2022.

[65] I. Skog, "Inertial and magnetic-field sensor arrays-capabilities and challenges," in *2018 IEEE SENSORS*, IEEE, 2018, pp. 1–4.

[66] M. H. Afzal, V. Renaudin, and G. Lachapelle, "Assessment of indoor magnetic field anomalies using multiple magnetometers," in *ION GNSS*, vol. 10, 2010, pp. 21–24.

[67] V. Renaudin, M. H. Afzal, and G. Lachapelle, "Complete triaxis magnetometer calibration in the magnetic domain," *Journal of sensors*, vol. 2010, 2010. DOI: 10.1155/2010/967245.

[68] M. H. Afzal, V. Renaudin, and G. Lachapelle, "Use of earth's magnetic field for mitigating gyroscope errors regardless of magnetic perturbation," *Sensors*, vol. 11, no. 12, pp. 11 390–11 414, 2011.

[69] V. Renaudin, M. H. Afzal, and G. Lachapelle, "New method for magnetometers based orientation estimation," in *IEEE/ION Position, Location and Navigation Symposium*, IEEE, 2010, pp. 348–356. DOI: 10.1109/PLANS.2010.5507301.

[70] G. Yin and L. Zhang, "Magnetic heading compensation method based on magnetic interferential signal inversion," *Sensors and Actuators A: Physical*, vol. 275, pp. 1–10, 2018.

[71] S. Guerrier, "Improving accuracy with multiple sensors: Study of redundant mems-imu/gps configurations," in *Proceedings of the 22nd international technical meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2009)*, 2009, pp. 3114–3121.

[72] X. Zhang, L. Xiao, C. Zou, L. Guo, Y. Jin, and G. Shi, "A novel calibration method of electronic compass based on multi-redundancy," in *2017 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, IEEE, 2017, pp. 156–161.

[73] I Colomina, M Giménez, J. Rosales, M Wis, A Gómez, and P Miguelsanz, "Redundant imus for precise trajectory determination," in *Proceedings of the 20th ISPRS Congress, Istanbul, Turkey*, Citeseer, vol. 1223, 2004, p. 17.

[74] S. Sukkarieh, P. Gibbens, B. Grocholsky, K. Willis, and H. F. Durrant-Whyte, "A low-cost, redundant inertial measurement unit for unmanned air vehicles," *The International Journal of Robotics Research*, vol. 19, no. 11, pp. 1089–1103, 2000.

[75] R. Giroux, S. Sukkarieh, and M. Bryson, "Implementation of a skewed-redundant low-cost ins in a fast-prototyping environment," in *Proceedings of the Institute of Navigation National Technical Meeting*, 2004.

[76] M. Jafari, "Optimal redundant sensor configuration for accuracy increasing in space inertial navigation system," *Aerospace Science and Technology*, vol. 47, pp. 467–472, 2015.

[77] Y. Yu and M. Cheng, "Comparative research of redundant strap down inertial navigation system based on different configuration schemes," in *AIP Conference Proceedings*, AIP Publishing, vol. 1967, 2018, p. 040 026.

[78] D. J. Allerton and H. Jia, "A review of multisensor fusion methodologies for aircraft navigation systems," *The Journal of Navigation*, vol. 58, no. 3, pp. 405–417, 2005. DOI: 10.1017/S0373463305003383.

[79] D. Unsal and K. Demirbas, "Estimation of deterministic and stochastic imu error parameters," in *Proceedings of the 2012 IEEE/ION Position, Location and Navigation Symposium*, IEEE, 2012, pp. 862–868.

[80] J. H. Wall, D. M. Bevly, *et al.*, "Characterization of various imu error sources and the effect on navigation performance," in *Proceedings of the 18th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2005)*, 2005, pp. 967–978.

[81] Q. Zhang, X. Niu, and C. Shi, "Impact assessment of various imu error sources on the relative accuracy of the gnss/ins systems," *IEEE Sensors Journal*, vol. 20, no. 9, pp. 5026–5038, 2020. DOI: 10.1109/JSEN.2020.2966379.

[82] D. Laidig and T. Seel, "Vqf: Highly accurate imu orientation estimation with bias estimation and magnetic disturbance rejection," *arXiv preprint arXiv:2203.17024*, 2022.

[83] L. Wang, H. Tang, T. Zhang, Q. Chen, J. Shi, and X. Niu, "Improving the navigation performance of the mems imu array by precise calibration," *IEEE Sensors Journal*, vol. 21, no. 22, pp. 26 050–26 058, 2021.

[84] J. Demkowicz and K. Bikonis, "Study of array of mems inertial measurement units under quasi-stationary and dynamic conditions," *Polish Maritime Research*, vol. 28, no. 3 (111), pp. 150–155, 2021.

[85] R. Siegwart, I. R. Nourbakhsh, and D. Scaramuzza, *Introduction to autonomous mobile robots*. MIT press, 2011.

[86] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[87] A. Karpathy, J. Johnson, and L. Fei-Fei, "Visualizing and understanding recurrent networks," *arXiv preprint arXiv:1506.02078*, 2015.

[88] D. Mandic and J. Chambers, *Recurrent neural networks for prediction: learning algorithms, architectures and stability*. Wiley, 2001, ISBN: 9780471495178.

[89] R. C. Staudemeyer and E. R. Morris, "Understanding lstm–a tutorial into long short-term memory recurrent neural networks," *arXiv preprint arXiv:1909.09586*, 2019.

[90] L. Medsker and L. C. Jain, *Recurrent neural networks: design and applications*. CRC press, 1999.

[91] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997. DOI: 10.1162/neco.1997.9.8.1735.

[92] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," *Neural computation*, vol. 12, no. 10, pp. 2451–2471, 2000. DOI: 10.1162/089976600300015015.

[93] F. A. Gers and J. Schmidhuber, "Recurrent nets that time and count," in *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*, IEEE, vol. 3, 2000, pp. 189–194.

[94] K. Yao, T. Cohn, K. Vylomova, K. Duh, and C. Dyer, "Depth-gated lstm," *arXiv preprint arXiv:1508.03790*, 2015.

[95] J. Koutnik, K. Greff, F. Gomez, and J. Schmidhuber, "A clockwork rnn," in *International Conference on Machine Learning*, PMLR, 2014, pp. 1863–1871.

[96] L. Nardi and C. Stachniss, "Actively improving robot navigation on different terrains using gaussian process mixture models," in *2019 International Conference on Robotics and Automation (ICRA)*, IEEE, 2019, pp. 4104–4110.

[97] A. A. Panchpor, S. Shue, and J. M. Conrad, "A survey of methods for mobile robot localization and mapping in dynamic indoor environments," in *2018 Conference on Signal Processing And Communication Engineering Systems (SPACES)*, IEEE, 2018, pp. 138–144.

[98] S. G. Tzafestas, "Mobile robot control and navigation: A global overview," *Journal of Intelligent & Robotic Systems*, vol. 91, no. 1, pp. 35–58, 2018.

[99] M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (slam) problem," *IEEE Transactions on robotics and automation*, vol. 17, no. 3, pp. 229–241, 2001.

[100] C. Cadena, L. Carlone, H. Carrillo, *et al.*, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Transactions on robotics*, vol. 32, no. 6, pp. 1309–1332, 2016.

[101] P. Pfreundschuh, H. F. Hendrikx, V. Reijgwart, R. Dubé, R. Siegwart, and A. Cramariuc, "Dynamic object aware lidar slam based on automatic generation of training data," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2021, pp. 11 641–11 647.

[102] A. Ahmadi, L. Nardi, N. Chebrolu, and C. Stachniss, "Visual servoing-based navigation for monitoring row-crop fields," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2020, pp. 4920–4926.

[103] R. Mur-Artal and J. D. Tardós, "Visual-inertial monocular slam with map reuse," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 796–803, 2017.

[104] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, "Orb-slam: A versatile and accurate monocular slam system," *IEEE transactions on robotics*, vol. 31, no. 5, pp. 1147–1163, 2015.

[105] T. Shan and B. Englot, "Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2018, pp. 4758–4765.

[106] 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 *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2020, pp. 5135–5142.

[107] R. Chatila and J. Laumond, "Position referencing and consistent world modeling for mobile robots," in *Proceedings. 1985 IEEE International Conference on Robotics and Automation*, vol. 2, Mar. 1985, pp. 138–145. DOI: 10.1109/ROBOT.1985.1087373.

[108] R. C. Smith and P. Cheeseman, "On the Representation and Estimation of Spatial Uncertainty," *The International Journal of Robotics Research*, vol. 5, no. 4, pp. 56–68, 1986. DOI: 10.1177/027836498600500404. [Online]. Available: https://doi.org/10.1177/027836498600500404.

[109] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part I," *IEEE Robotics Automation Magazine*, vol. 13, no. 2, pp. 99–110, Jun. 2006, ISSN: 1558-223X. DOI: 10.1109/MRA.2006.1638022.

[110] C. Cadena, L. Carlone, H. Carrillo, *et al.*, "Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1309–1332, Dec. 2016, ISSN: 1941-0468. DOI: 10.1109/TRO.2016.2624754.

[111] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): part II," *IEEE Robotics Automation Magazine*, vol. 13, no. 3, pp. 108–117, Sep. 2006, ISSN: 1558-223X. DOI: 10.1109/MRA.2006.1678144.

[112] C. Stachniss, J. J. Leonard, and S. Thrun, "Simultaneous Localization and Mapping," in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Cham: Springer International Publishing, 2016, ch. 46, pp. 1153–1176, ISBN: 978-3-319-32552-1. DOI: 10.1007/978-3-319-32552-1_46. [Online]. Available: https://doi.org/10.1007/978-3-319-32552-1_46.

[113] G. Dissanayake, S. Huang, Z. Wang, and R. Ranasinghe, "A review of recent developments in Simultaneous Localization and Mapping," in *2011 6th International Conference on Industrial and Information Systems*, Aug. 2011, pp. 477–482. DOI: 10.1109/ICIINFS.2011.6038117.

[114] G. Grisetti, R. Kümmerle, C. Stachniss, and W. Burgard, "A Tutorial on Graph-Based SLAM," *IEEE Intelligent Transportation Systems Magazine*, vol. 2, no. 4, pp. 31–43, 2010, ISSN: 1941-1197. DOI: 10.1109/MITS.2010.939925.

[115] S. Saeedi, M. Trentini, M. Seto, and H. Li, "Multiple-robot simultaneous localization and mapping: A review," *Journal of Field Robotics*, vol. 33, no. 1, pp. 3–46, 2016. DOI: 10.1002/rob.21620.

[116] G. Bresson, Z. Alsayed, L. Yu, and S. Glaser, "Simultaneous Localization and Mapping: A Survey of Current Trends in Autonomous Driving," *IEEE Transactions on Intelligent Vehicles*, vol. 2, no. 3, pp. 194–220, Sep. 2017, ISSN: 2379-8858. DOI: 10.1109/TIV.2017.2749181.

[117] J. Fuentes-Pacheco, J. Ruiz-Ascencio, and J. M. Rendón-Mancha, "Visual simultaneous localization and mapping: A survey," *Artificial Intelligence Review*, vol. 43, no. 1, pp. 55–81, 2015, ISSN: 1573-7462. DOI: 10.1007/s10462-012-9365-8. [Online]. Available: https://doi.org/10.1007/s10462-012-9365-8.

[118] M. R. U. Saputra, A. Markham, and N. Trigoni, "Visual SLAM and Structure from Motion in Dynamic Environments: A Survey," *ACM Computing Surveys*, vol. 51, no. 2, Feb. 2018, ISSN: 0360-0300. DOI: 10.1145/3177853. [Online]. Available: https://doi.org/10.1145/3177853.

[119] B. Huang, J. Zhao, and J. Liu, *A Survey of Simultaneous Localization and Mapping with an Envision in 6G Wireless Networks*, Feb. 2020. arXiv: 1909.05214 [cs.RO].

[120] D. Van Opdenbosch and E. Steinbach, "Collaborative visual slam using compressed feature exchange," *IEEE Robotics and Automation Letters*, vol. 4, no. 1, pp. 57–64, 2018.

[121] A. Garcea, J. Zhu, D. Van Opdenbosch, and E. Steinbach, "Robust map alignment for cooperative visual slam," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, IEEE, 2018, pp. 4083–4087.

[122] Y. Zhao, W. Ye, and P. A. Vela, "Low-latency visual slam with appearance-enhanced local map building," in *2019 International Conference on Robotics and Automation (ICRA)*, IEEE, 2019, pp. 8213–8219.

[123] Q. Zou, Q. Sun, L. Chen, B. Nie, and Q. Li, "A comparative analysis of lidar slam-based indoor navigation for autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[124] C. Qu, S. S. Shivakumar, W. Liu, and C. J. Taylor, "Llol: Low-latency odometry for spinning lidars," *arXiv preprint arXiv:2110.01725*, 2021. DOI: 10.48550/arXiv.2110.01725.

[125] A. Al-Nuaimi, W. Lopes, P. Zeller, A. Garcea, C. Lopes, and E. Steinbach, "Analyzing lidar scan skewing and its impact on scan matching," in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, IEEE, 2016, pp. 1–8.

[126] N. Zarrabi, R. Fesharakifard, and M. B. Menhaj, "Robot localization performance using different slam approaches in a homogeneous indoor environment," in *2019 7th International Conference on Robotics and Mechatronics (ICRoM)*, IEEE, 2019, pp. 338–344.

[127] Z. Liu, H. Chen, H. Di, *et al.*, "Real-time 6d lidar slam in large scale natural terrains for ugv," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 2018, pp. 662–667.

[128] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, Feb. 1992, ISSN: 1939-3539. DOI: 10.1109/34.121791.

[129] Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *International journal of computer vision*, vol. 13, no. 2, pp. 119–152, 1994. DOI: 10.1007/bf01427149.

[130] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, May 2001, pp. 145–152. DOI: 10.1109/IM.2001.924423.

[131] F. Pomerleau, F. Colas, and R. Siegwart, "A Review of Point Cloud Registration Algorithms for Mobile Robotics," *Foundations and Trends in Robotics*, vol. 4, no. 1, pp. 1–104, 2015. DOI: 10.1561/2300000035. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01178661.

[132] A. Al-Nuaimi, "Methods of Point Cloud Alignment with Applications to 3D Indoor Mapping and Localization," Dissertation, Technische Universität München, München, 2017. [Online]. Available: https://mediatum.ub.tum.de/1327421.

[133] J. Zhang, "Online Lidar and Vision based Ego-motion Estimation and Mapping," Ph.D. dissertation, Carnegie Mellon University, Pittsburgh, PA, Mar. 2017. [Online]. Available: https://www.ri.cmu.edu/publications/online-lidar-vision-based-ego-motion-estimation-mapping/.

[134] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, and M. Levoy, "Geometrically stable sampling for the ICP algorithm," in *Fourth International Conference on 3-D Digital Imaging and Modeling, 2003. 3DIM 2003. Proceedings.*, IEEE, Oct. 2003, pp. 260–267. DOI: 10.1109/IM.2003.1240258.

[135] J.-E. Deschaud, "IMLS-SLAM: scan-to-model matching based on 3D data," in *ICRA - International Conference on Robotics and Automation*, ser. 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, May 2018. DOI: 10.1109/icra.2018.8460653. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01959570.

[136] M. Bosse and R. Zlot, "Continuous 3D scan-matching with a spinning 2D laser," in *2009 IEEE International Conference on Robotics and Automation*, May 2009, pp. 4312–4319. DOI: 10.1109/ROBOT.2009.5152851.

[137] M. Schadler, J. Stückler, and S. Behnke, "Multi-resolution surfel mapping and real-time pose tracking using a continuously rotating 2D laser scanner," in *2013 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, Oct. 2013, pp. 1–6. DOI: 10.1109/SSRR.2013.6719373.

[138] J. Behley and C. Stachniss, "Efficient Surfel-Based SLAM using 3D Laser Range Data in Urban Environments.," in *Robotics: Science and Systems*, 2018. DOI: 10.15607/rss.2018.xiv.016.

[139] C. Park, S. Kim, P. Moghadam, C. Fookes, and S. Sridharan, "Probabilistic Surfel Fusion for Dense LiDAR Mapping," in *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, Oct. 2017, pp. 2418–2426. DOI: 10.1109/ICCVW.2017.285.

[140] D. Droeschel, M. Schwarz, and S. Behnke, "Continuous mapping and localization for autonomous navigation in rough terrain using a 3D laser scanner," *Robotics and Autonomous Systems*, vol. 88, pp. 104–115, 2017, ISSN: 0921-8890. DOI: https://doi.org/10.1016/j.robot.2016.10.017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0921889015303110.

[141] J. Zhang and S. Singh, "LOAM: Lidar Odometry and Mapping in Real-time," in *Proceedings of Robotics: Science and Systems*, Berkeley, USA, Jul. 2014. DOI: 10.15607/RSS.2014.X.007.

[142] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2012, pp. 3354–3361. DOI: 10.1109/CVPR.2012.6248074.

[143] J. Zhang and S. Singh, "Visual-lidar odometry and mapping: Low-drift, robust, and fast," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, May 2015, pp. 2174–2181. DOI: 10.1109/ICRA.2015.7139486.

[144] T. Shan and B. Englot, "LeGO-LOAM: Lightweight and Ground-Optimized Lidar Odometry and Mapping on Variable Terrain," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct. 2018, pp. 4758–4765. DOI: 10.1109/IROS.2018.8594299.

[145] M. Yan, J. Wang, J. Li, and C. Zhang, "Loose coupling visual-lidar odometry by combining VISO2 and LOAM," in *2017 36th Chinese Control Conference (CCC)*, Jul. 2017, pp. 6841–6846. DOI: 10.23919/ChiCC.2017.8028435.

[146] X. Ji, L. Zuo, C. Zhang, and Y. Liu, "LLOAM: LiDAR Odometry and Mapping with Loop-closure Detection Based Correction," in *2019 IEEE International Conference on Mechatronics and Automation (ICMA)*, Aug. 2019, pp. 2475–2480. DOI: 10.1109/ICMA.2019.8816388.

[147] Z. Youji, C. Qijun, Z. Hao, L. Dairong, and W. Penghao, "A SLAM Method Based on LOAM for Ground Vehicles in the Flat Ground," in *2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS)*, May 2019, pp. 546–551. DOI: 10.1109/ICPHYS.2019.8780378.

[148] X. Liu, L. Zhang, S. Qin, D. Tian, S. Ouyang, and C. Chen, "Optimized LOAM Using Ground Plane Constraints and SegMatch-Based Loop Detection," *Sensors*, vol. 19, no. 24, p. 5419, Dec. 2019, ISSN: 1424-8220. DOI: 10.3390/s19245419. [Online]. Available: http://dx.doi.org/10.3390/s19245419.

[149] J. Qian, X. Mai, and X. Yuwen, "Real-time Power Line Safety Distance Detection System Based on LOAM Slam," in *2018 Chinese Automation Congress (CAC)*, Nov. 2018, pp. 3204–3208. DOI: 10.1109/CAC.2018.8623168.

[150] C. Beokhaimook, K. Ohno, H. Nishinoman, T. Hoshi, and S. Tadokoro, "Development of Cyber-enhanced Rescue Canine Suit for 3D SLAM," *The Proceedings of JSME annual Conference on Robotics and Mechatronics (Robomec)*, vol. 2019, 2A2–E07, Dec. 2019. DOI: 10.1299/jsmermd.2019.2A2-E07.

[151] T. Raj, F. H. Hashim, A. B. Huddin, M. F. Ibrahim, and A. Hussain, "A survey on lidar scanning mechanisms," *Electronics*, vol. 9, no. 5, p. 741, 2020. DOI: 10.3390/electronics9050741.

[152] Y. Li and E. B. Olson, "Structure tensors for general purpose LIDAR feature extraction," in *2011 IEEE International Conference on Robotics and Automation*, May 2011, pp. 1869–1874. DOI: 10.1109/ICRA.2011.5979567.

[153] M. Muglikar, Z. Zhang, and D. Scaramuzza, "Voxel map for visual slam," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2020, pp. 4181–4187. DOI: 10.1109/ICRA40945.2020.9197357.

[154] X.-F. Han, J. S. Jin, M.-J. Wang, W. Jiang, L. Gao, and L. Xiao, "A review of algorithms for filtering the 3d point cloud," *Signal Processing: Image Communication*, vol. 57, pp. 103–112, 2017. DOI: 10.1016/j.image.2017.05.009.

[155] J. Ryde and H. Hu, "3d mapping with multi-resolution occupied voxel lists," *Autonomous Robots*, vol. 28, no. 2, pp. 169–185, 2010. DOI: 10.1007/s10514-009-9158-3.

[156] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "Octomap: An efficient probabilistic 3d mapping framework based on octrees," *Autonomous robots*, vol. 34, no. 3, pp. 189–206, 2013. DOI: 10.1007/s10514-012-9321-0.

[157] P. Ram and K. Sinha, "Revisiting kd-tree for nearest neighbor search," in *Proceedings of the 25th acm sigkdd international conference on knowledge discovery & data mining*, 2019, pp. 1378–1388. DOI: 10.1145/3292500.3330875.

[158] F. Valenti, D. Giaquinto, L Musto, A Zinelli, M. Bertozzi, and A. Broggi, "Enabling computer vision-based autonomous navigation for unmanned aerial vehicles in cluttered gps-denied environments," in *International Conference on Intelligent Transportation Systems*, 2018.

[159] J. Siva and C. Poellabauer, "Robot and drone localization in gps-denied areas," in *Mission-Oriented Sensor Networks and Systems: Art and Science*, Springer, 2019, pp. 597–631.

[160] Y. D. Yasuda, L. E. G. Martins, and F. A. Cappabianco, "Autonomous visual navigation for mobile robots: A systematic literature review," *ACM Computing Surveys (CSUR)*, vol. 53, no. 1, pp. 1–34, 2020. DOI: 10.1145/3368961.

[161] X. Zhou, Z. Yi, Y. Liu, K. Huang, and H. Huang, "Survey on path and view planning for uavs," *Virtual Reality & Intelligent Hardware*, vol. 2, no. 1, pp. 56–69, 2020.

[162] S. Jung, S. Hwang, H. Shin, and D. H. Shim, "Perception, Guidance, and Navigation for Indoor Autonomous Drone Racing Using Deep Learning," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2539–2544, 2018, ISSN: 2377-3766. DOI: 10.1109/LRA.2018.2808368.

[163] R. S. Dimitrova, M. Gehrig, D. Brescianini, and D. Scaramuzza, *Towards Low-Latency High-Bandwidth Control of Quadrotors using Event Cameras*, 2019. arXiv: 1911.04553 [cs.RO].

[164] R. Pérez-Alcocer and J. Moreno-Valenzuela, "A novel lyapunov-based trajectory tracking controller for a quadrotor: Experimental analysis by using two motion tasks," *Mechatronics*, vol. 61, pp. 58–68, 2019. DOI: 10.1016/j.mechatronics.2019.05.006.

[165] M. Velas, M. Spanel, T. Sleziak, J. Habrovec, and A. Herout, "Indoor and outdoor backpack mapping with calibrated pair of velodyne lidars," *Sensors*, vol. 19, no. 18, p. 3944, 2019. DOI: 10.3390/s19183944.

[166] J. Lin and F. Zhang, "Loam_livox: A fast, robust, high-precision lidar odometry and mapping package for lidars of small fov," *arXiv preprint arXiv:1909.06700*, 2019. DOI: 10.1109/ICRA40945.2020.9197440.

[167] ——, "A fast, complete, point cloud based loop closure for lidar odometry and mapping," *arXiv preprint arXiv:1909.11811*, 2019. DOI: 10.48550/arXiv.1909.11811.

[168] H. Qin, Y. Bi, K. Z. Ang, *et al.*, "A stereo and rotating laser framework for uav navigation in gps denied environment," in *Industrial Electronics*, IEEE, 2016. DOI: 10.1109/IECON.2016.7793246.

[169] W. Zhen and S. Scherer, "A unified 3d mapping framework using a 3d or 2d lidar," *arXiv preprint arXiv:1810.12515*, 2018. DOI: 10.1007/978-3-030-33950-0_60.

[170] H. Qin, Z. Meng, W. Meng, *et al.*, "Autonomous exploration and mapping system using heterogeneous uavs and ugvs in gps-denied environments," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1339–1350, 2019.

[171] R. Voges, C. S. Wieghardt, and B. Wagner, "Timestamp offset determination between an actuated laser scanner and its corresponding motor," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, p. 99, 2017. DOI: 10.15488/3396.

[172] J. Zhang and S. Singh, "Low-drift and real-time lidar odometry and mapping," *Autonomous Robots*, vol. 41, no. 2, pp. 401–416, 2017. DOI: 10.1007/s10514-016-9548-2.

[173] L. Kaul, R. Zlot, and M. Bosse, "Continuous-time three-dimensional mapping for micro aerial vehicles with a passively actuated rotating laser scanner," *Journal of Field Robotics*, vol. 33, no. 1, 2016. DOI: 10.1002/rob.21614.

[174] S. Anderson and T. D. Barfoot, "Towards relative continuous-time slam," in *2013 IEEE International Conference on Robotics and Automation*, IEEE, 2013, pp. 1033–1040. DOI: 10.1109/ICRA.2013.6630700.

[175] M. Bosse, R. Zlot, and P. Flick, "Zebedee: Design of a spring-mounted 3-d range sensor with application to mobile mapping," *IEEE Transactions on Robotics*, vol. 28, no. 5, pp. 1104–1119, 2012. DOI: 10.1109/TRO.2012.2200990.

[176] J. W. Anderson and G. M. Clayton, "Lissajous-like scan pattern for a gimballed lidar," in *2014 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, IEEE, 2014, pp. 1171–1176. DOI: 10.1109/AIM.2014.6878240.

[177] M. Benson, J. Nikolaidis, and G. M. Clayton, "Lissajous-like scan pattern for a nodding multi-beam lidar," in *ASME 2018 Dynamic Systems and Control Conference*, American Society of Mechanical Engineers Digital Collection, 2018. DOI: 10.1115/DSCC2018-9169.

[178] D. Droeschel, M. Nieuwenhuisen, M. Beul, D. Holz, J. Stückler, and S. Behnke, "Multilayered mapping and navigation for autonomous micro aerial vehicles," *Journal of Field Robotics*, vol. 33, no. 4, 2016. DOI: 10.1002/rob.21603.

[179] H. Nakagomi, Y. Fuse, H. Hosaka, *et al.*, "3d scan matching for mobile robot localization over rough terrain," *Electrical Engineering in Japan*, vol. 209, no. 3-4, 2019. DOI: 10.1002/eej.23250.

[180] T. Yoshida, K. Irie, E. Koyanagi, and M. Tomono, "3D laser scanner with gazing ability," in *2011 IEEE International Conference on Robotics and Automation*, May 2011, pp. 3098–3103. DOI: 10.1109/ICRA.2011.5980385.

[181] H. Dong and T. D. Barfoot, "Lighting-invariant visual odometry using lidar intensity imagery and pose interpolation," in *Field and Service Robotics*, Springer, 2014, pp. 327–342. DOI: 10.1007/978-3-642-40686-7_22.

List of Figures