RESSALVA

Atendendo solicitação do(a) autor(a), o texto completo desta tese será disponibilizado somente a partir de 29/05/2027.

UNESP – UNIVERSIDADE ESTADUAL PAULISTA
CAMPUS DE PRESIDENTE PRUDENTE
FACULDADE DE CIÊNCIAS E TECNOLOGIA
PÓS-GRADUAÇÃO EM CIÊNCIAS CARTOGRÁFICAS

THAISA ALINE CORREIA GARCIA

REAL-TIME VISUAL POSITIONING WITH OMNIDIRECTIONAL IMAGES

Presidente Prudente/SP

2025

UNESP – UNIVERSIDADE ESTADUAL PAULISTA
CAMPUS DE PRESIDENTE PRUDENTE
FACULDADE DE CIÊNCIAS E TECNOLOGIA
PÓS-GRADUAÇÃO EM CIÊNCIAS CARTOGRÁFICAS

THAISA ALINE CORREIA GARCIA

REAL-TIME VISUAL POSITIONING WITH OMNIDIRECTIONAL IMAGES

A Thesis submitted to the Graduate program in Cartographic Sciences (PPGCC) at the Faculty of Science and Technology, São Paulo State University, in partial fulfilment of the requirements for the degree of Doctor of Cartographic Sciences.

Supervisor: Prof. Dr. Antonio M. G. Tommaselli
Co- Supervisor: Dr. Mariana Batista Campos

Presidente Prudente/SP
2025

G216r

Garcia, Thaisa Aline Correia
Real-time visual positioning with omnidirectional images / Thaisa Aline Correia Garcia. -- Presidente Prudente, 2026
136 p. : il., tabs., fotos, mapas

Tese (doutorado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Tecnologia, Presidente Prudente
Orientador: Antonio Maria Garcia Tommaselli
Coorientadora: Mariana Batista Campos

1. Visual Simultaneous Localization and Mapping. 2. Fotogrametria. 3. Sistema omnidirecional. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Dados fornecidos pelo autor(a).

IMPACTO POTENCIAL DESTA PESQUISA

Esta pesquisa desenvolve um sistema de posicionamento visual com câmeras onidirecionais, aplicável a plataformas autônomas. Modelos geométricos rigorosos e filtragem de pontos-chave por aprendizado de máquina aprimoram a precisão e desempenho em tempo real, viabilizando aplicações em agricultura de precisão, robótica e monitoramento ambiental.

POTENTIAL IMPACT OF THIS RESEARCH

This research develops a visual positioning system using omnidirectional cameras, applicable to autonomous platforms. Rigorous geometric models and machine-learning-based keypoint filtering improve accuracy and real-time performance, enabling applications in precision agriculture, mobile robotics and environmental monitoring.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Presidente Prudente

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: **REAL-TIME VISUAL POSITIONING WITH OMNIDIRECTIONAL IMAGES**

AUTORA: THAÍSA ALINE CORREIA GARCIA
ORIENTADOR: ANTONIO MARIA GARCIA TOMMASELLI
COORIENTADORA: MARIANA BATISTA CAMPOS

Aprovada como parte das exigências para obtenção do Título de Doutora em Ciências Cartográficas, área: Aquisição, Análise e Representação de Informações Espaciais pela Comissão Examinadora:

Prof. Dr. ANTONIO MARIA GARCIA TOMMASELLI (Participaçao Presencial)
Departamento de Cartografia / FCT UNESP Presidente Prudente

Prof. Dr. MAURICIO GALO (Participaçao Presencial)
Departamento de Cartografia / Faculdade de Ciencias e Tecnologia de Presidente Prudente - FCT/Unesp

Prof. Dr. RENATO CÉSAR DOS SANTOS (Participaçao Presencial)
Faculdade de Ciências e Tecnologia de Presidente Prudente

Prof. Dr. DANIEL RODRIGUES DOS SANTOS (Participaçao Virtual)
Seção de Engenharia Cartográfica / Instituto Militar de Engenharia

Prof. Dr. HENRIQUE CANDIDO DE OLIVEIRA (Participaçao Virtual)
Departamento de Infraestrutura e Ambiente / Universidade Estadual de Campinas

Presidente Prudente, 29 de maio de 2025

To my God, who has sustained me thus far, and to my beloved Marcos, Liz and Levi, for their unwavering support, love, and understanding throughout this journey.

ACKNOWLEDGMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

This study was financed, in part, by the São Paulo Research Foundation (FAPESP), Brasil. Process Number 2021/06029-7. The opinions, hypotheses and conclusions or recommendations expressed in this material are the responsibility of the authors and do not necessarily reflect the views of FAPESP.

The authors also thank the National Council for Scientific and Technological Development (CNPq) for the support (grants n. 141550/2020-1 and 303670/2018-5).

I would like to express my gratitude to FCT-Unesp (Faculty of Science and Technology - São Paulo State University) and the PPGCC (Graduate program in Cartographic Sciences) for the opportunity, support, and infrastructure that made this research possible.

I am deeply grateful to my advisor, Prof. Dr. Antonio Maria Garcia Tommaselli, and my co-advisor, Dr. Mariana Batista Campos, for their guidance, support, understanding, and teaching throughout this journey. It is an honour to work with you.

Special thanks to my friends Marcelo Pacheco and Thiago Tiedtke for introducing me to the world of Photogrammetry, for their constant support and encouragement, and for their generosity in sharing their expertise.

A thank you to all my colleagues at PPGCC for their willingness to share knowledge and support each other through collaboration, especially to Letícia Ferrari Castanheiro, who assisted me throughout the entire data capture and processing process.

RESUMO

A estimação do posicionamento e da orientação de uma plataforma móvel em tempo real com a reconstrução simultânea de um mapa do ambiente é um problema fundamental em diversas áreas, como robótica móvel, desenvolvimento de veículos autônomos, realidade aumentada e mapeamento 3D. Algumas abordagens enfrentam dificuldades em ambientes externos onde sinais GNSS (*Global Navigation Satellite System*) são pouco confiáveis ou inexistentes, especialmente em florestas densas ou corredores urbanos longos. Nesses cenários, técnicas de navegação com sensores ópticos, como o vSLAM (*visual Simultaneous Localization and Mapping*) oferecem uma alternativa, utilizando imagens para estimar a trajetória e a nuvem de pontos 3D. No entanto, essas soluções são afetadas por erros causados por mudanças na iluminação, oclusões e limitações na sobreposição de imagens. As câmaras *fisheye*, devido ao seu amplo campo de visão, demonstram potencial para mitigar algumas dessas limitações, capturando mais informações contextuais a cada imagem. Contudo, o uso dessas lentes apresenta desafios de modelagem matemática, uma vez que o modelo de colinearidade padrão não é adequado. O presente trabalho propõe uma solução de posicionamento em tempo real para uma plataforma móvel (mochila), utilizando um sistema omnidirecional dual de câmaras com lentes *fisheye* hiperhemisféricas (Ricoh Theta S), com um rigoroso modelo de projeção equisólido para melhorar a precisão da estimativa de trajetória. A abordagem de mapeamento móvel terrestre utiliza processamento paralelo com os recursos de CPU e GPU para alcançar desempenho em tempo real, mesmo em ambientes complexos. A abordagem desenvolvida adaptou o software ORB-SLAM para geometrias fisheye e avaliou seu desempenho na plataforma Nvidia Jetson TX2, com foco em aplicações de navegação autônoma, em que retornar ao ponto de partida não é viável. A solução proposta aborda desafios-chave no processamento em tempo real e na consistência da trajetória, com foco em ambientes onde o sinal GNSS está degradado ou não disponível. Os resultados evidenciaram melhorias consistentes na acurácia da trajetória com a abordagem proposta: ao final de uma trajetória de 140 m, o modelo equisólido apresentou RMSE planimétrico em E de 3,94 m, frente a 40,86 m do modelo genérico EUCM.

Palavras-chaves: vSLAM; Fotogrametria; sistema omnidirecional; tempo real; ORB-SLAM.

ABSTRACT

The estimation of the position and orientation of a mobile platform in real-time while simultaneously reconstructing a map of the environment is a fundamental challenge in several areas, such as mobile robotics, autonomous vehicles, augmented reality and 3D mapping. Some approaches face challenges in outdoor environments where GNSS signals are unreliable or unavailable, particularly in dense forests or long corridors. In such scenarios, optical sensor-based navigation techniques, such as visual Simultaneous Localization and Mapping (vSLAM), provide an alternative by leveraging images to estimate trajectory and generate a 3D point cloud. However, these methods are prone to errors caused by lighting changes, occlusions, and limited image overlap. Fisheye cameras, with their wide FoV, show promise in mitigating some of these limitations by capturing more contextual information in each frame. Nonetheless, their use introduces challenges in mathematical modelling, as the standard collinearity model is inadequate. This work proposes a real-time positioning solution for a mobile backpack platform, employing a dual omnidirectional camera system featuring hyperhemispherical fisheye lenses (Ricoh Theta S), with a rigorous equisolid projection model to improve trajectory estimation accuracy. The terrestrial mobile mapping approach utilises parallel processing on CPU and GPU resources to achieve real-time performance, even in challenging environments. The developed approach adapts the ORB-SLAM software to fisheye geometries and evaluates its performance on the Nvidia Jetson TX2 platform, targeting autonomous navigation applications, where returning to the starting point is not feasible. The proposed solution addresses key challenges in real-time processing and trajectory consistency, with a focus on environments where GNSS signals are degraded or unavailable. The results evidenced consistent improvements in trajectory accuracy with the proposed approach: at the end of a 140 m trajectory, the equisolid model yielded a planimetric RMSE in E of 3.94 m, compared to 40.86 m for the generic EUCM model.

Keywords: vSLAM; Photogrammetry; omnidirectional system; real time; ORB-SLAM.

LIST OF SYMBOLS AND ABBREVIATION

2D	Two-dimensional
3D	Three-dimensional
BA	Bundle Adjustment
BRIEF	Binary Robust Independent Elementary Features
CMOS	Complementary Metal-Oxide-Semiconductor
CNN	Convolutional Neural Networks
CPU	Central Processing Unit
EKF	Extended Kalman Filter
EOP	Exterior Orientation Parameters
EUCM	Enhanced Unified Camera Model
FAST	Features from Accelerated Segment Test
FBM	Feature-Based Matching
FoV	Field of View
GCP	Ground Control Points
GNSS	Global Navigation Satellite System
GPU	Graphics Processing Unit
LiDAR	Light Detection And Ranging
LSM	Least Squares Method
IMU	Inertial Measurement Unit
IOP	Interior Orientation Parameters
ORB	Oriented FAST and rotated BRIEF
PnP	Perspective-n-Points
RADAR	Radio Detection and Ranging
RANSAC	Random Sample Consensus
RF	Random Forest
ROP	Relative Orientation Parameters
RTK	Real Time Kinematic
SIFT	Scale-Invariant Feature Transform
SLAM	Simultaneous Localisation and Mapping
SfM	Structure from Motion
SVD	Singular Value Decomposition

UTM	Universal Transverse Mercator
vSLAM	Visual Simultaneous Localization and Mapping

LIST OF FIGURES

LIST OF TABLES

SUMMARY

CHAPTER 1
INTRODUCTION

1.1 OVERVIEW

The classic problem of sequentially estimating an agent's movement in real time, while maintaining a consistent trajectory based on a 3D point map of the environment, has gained significant popularity in the last decade. This growth is mainly due to the introduction of various sensors and the increasing computational power of embedded platforms (Cadena et al., 2016). Several techniques address this problem, including Simultaneous Localization and Mapping (SLAM). SLAM is a subfield of mobile robotic research that relies on sensors such as LiDAR (Light Detection And Ranging) and/or optical cameras. An overview of the evolution of the main methodologies in the literature can be found in Durrant-Whyte and Bailey (2006) and Alsadik and Karam (2021).

Real-time sequential positioning using optical sensors began to emerge in the early 1980s. This problem was mainly studied for close-range mobile platforms, applying well-known photogrammetric techniques (Haggren, 1986; Gruen, 1988; Durrant-Whyte and Bailey, 2006; Scaramuzza and Fraundorfer, 2011; Cadena et al., 2016). Moravec (1980) made a significant contribution to the field by describing a system for a robot to navigate autonomously in an unknown environment and to create a map of its surroundings using stereovision. Moravec (1980) also described one of the first interest point detectors, considered a precursor to the one proposed by Forstner (1986). The appearance of corner detectors and advancements in technology have driven studies and research in this area over the past thirty years, resulting in significant progress in solutions applicable to a wide range of fields, including autonomous navigation systems, robotics, industry, agriculture, and many others (Scaramuzza, Fraundorfer and Siegwart, 2009; Cadena et al., 2016).

When the SLAM approach relies exclusively on sequential images captured by optical cameras, it is referred to as Visual Simultaneous Localization and Mapping (vSLAM), which is widely used in indoor solutions where Global Navigation Satellite System (GNSS) signal obstructions occurs (Scaramuzza, Fraundorfer and Siegwart, 2009). The availability of low-cost, lightweight optical sensors and high-processing-power embedded platforms has driven the development of vSLAM approaches in recent years (Gao and Zhang, 2021). Among the key approaches are Mono-SLAM (Davison et al., 2007), Parallel Tracking and Mapping for Small AR Workspaces (PTAM) (Klein and Murray, 2007), Dense Tracking and Mapping in Real-Time (DTAM) (Newcombe, Lovegrove and Davison, 2011), Fast Semi-

Direct Monocular Visual Odometry (SVO) (Forster, Pizzoli and Scaramuzza, 2014), Large-Scale Direct Monocular SLAM (LSD-SLAM) (Engel, Schõps and Cremers, 2014), ORB-SLAM (Mur-Artal, Montiel and Tardos, 2015), and Direct Sparse Odometry with Fisheye Cameras (DSO) (Matsuki et al., 2018). ORB-SLAM is regarded as one of the state-of-the-art vSLAM solutions that uses Oriented FAST and rotated BRIEF (ORB) as keypoints detector. It is based on a modular, multi-threaded architecture composed of the main methods and approaches of vSLAM, and was initially developed for perspective images (Mur-Artal, Montiel and Tardos, 2015). Simultaneously to those developments, cameras equipped with wide field-of-view (FoV) lenses, such as fisheye lenses, have become increasingly popular on vSLAM applications, but their use requires adaptations to previously mentioned methods, such as ORB-SLAM.

Fisheye lens cameras, equipped with hemispherical lenses (FoV ~ 180°) or hyperhemispherical lenses (FoV > 180°), can capture a broader view of the scene in a single image, encompassing more points and contextual information while maintaining lower complexity compared to conventional multi-camera perspective setups (Schneider and Forstner, 2013; Caruso, Engel and Cremers, 2015; Zhang et al., 2016; Matsuki et al., 2018; Wang et al., 2018; Liu et al., 2019; Campos et al., 2020). Some works proposed extending the use of ORB-SLAM with wider FoV images (Liu et al., 2019; Campos et al., 2020; Zhang; Huang, 2021; Zhao et al., 2022), and multi-camera omnidirectional systems (Zhao et al., 2022), especially by changing the original perspective mathematical model (pinhole model) for generics or unified mathematical models, such as Enhanced Unified Camera Model (EUCM) (Khomutenko, Garcia and Martinet, 2016). This is necessary because the collinearity equations, combined with distortion models, do not model the internal fisheye camera geometry, in which incident rays are deflected after being refracted by the camera lenses (Schneider, Schwalbe and Maas, 2009). Therefore, other projection models need to be considered. As mentioned, most of fisheye vSLAM solutions use generic or unified mathematical models for fisheye cameras, enabling the application of a simplified model without the need for a physical interpretation of the image formation geometry (Kannala and Brandt, 2006; Courbon et al., 2007; Khomutenko, Garcia and Martinet, 2016; Usenko, Demmel and Cremers, 2018). However, the accuracy of the vSLAM solution can be affected by not considering rigorously the physical parameters, and generic models may not fully account for all errors by failing to mathematically describe the image formation process. This leads to the propagation of these errors in a sequential solution (Schneider, Schwalbe and

Maas, 2009; Garcia et al., 2023; Tommaselli et al., 2023). Therefore, further studies are still needed, especially when using multiple cameras to achieve a 360° field of view.

Multi-camera omnidirectional systems tend to be more effective in complex environments compared to monocular systems, as they process a wider FoV with greater level of details (Li, Wang and Li, 2018; Zhang and Huang, 2021; Song, Liu and Dai, 2024; Xie et al., 2024; Yang et al., 2024). Currently, most omnidirectional vSLAM methods use multiple cameras to track more interest points, enhancing both accuracy and robustness. Optical omnidirectional systems typically provide a 360° horizontal FoV or a full-spherical panorama. These multi-camera systems can be designed with conventional lenses, such as perspective lenses, or with non-conventional optical systems, like fisheye lenses. However, the use of 360° images in photogrammetry still requires further research in rigorous camera modelling, as most of these systems were not designed for metric purposes (Campos, 2019). Campos et al. (2018b) proposed a post-processed solution for an embedded mobile platform, consisting of a 360° omnidirectional camera, a single-frequency GNSS receiver, and a low-cost (Inertial Measurement Unit) IMU. In this work it was developed an incremental Bundle Adjustment (BA) strategy to estimate the Exterior Orientation Parameters (EOP) of the sensors and 3D points, based on data collected by the system, using a rigorous spherical mathematical model aligned with the geometry of dual polydioptric systems based on fisheye lenses.

Despite the significant advances mentioned, vSLAM solutions for fisheye lenses and omnidirectional dual-camera geometries have not been fully explored, particularly considering the additional challenge of real-time processing using a rigorous mathematical model, which requires greater computational resources. vSLAM frameworks, such as ORB-SLAM, require substantial memory and processing power. High computational resources are required to run processes to mitigate the errors, which can result in divergence of the solution, such as outlier detection and false correspondences removal while estimating sensor orientations and 3D coordinates (Mur-Artal, Montiel and Tardos, 2015). This demand can limit their real-time performance on mobile platforms with limited Central Processing Unit (CPU) and Graphics Processing Unit (GPU) resources, making computational resource optimisation challenging. Although ORB-SLAM can operate in real-time on conventional computers, it does not achieve the same level of performance on mobile platforms due to the complexity of the operations, particularly in the steps of feature extraction and matching. The data processing complexity increases in challenging environments involving lengthy, linear outdoor trajectories, such as dense forests or perennial crop fields, characterized by tall trees and dense canopy coverage. In these areas, reliance on external positioning information is either unavailable or unreliable, due to the limitations of GNSS signal in these settings or the high noise level and cost associated with IMU sensor data (Caruso, Engel and Cremers, 2015; Zhao et al., 2022). Furthermore, in addition to the challenges faced by purely visual solutions, such as variations in lighting conditions, shadows, occlusions, moving objects, and homogeneous features, unsuitable mathematical modelling can introduce additional sources of error that may lead to divergence, mainly in such challenging environments (Scaramuzza, Fraundorfer and Siegwart, 2009; Matsuki et al., 2018; Tommaselli et al., 2023). Therefore, initiatives aimed at reducing errors for real-time applications are still necessary, especially those focused on exploring parallelization methodologies to improve data processing efficiency in challenging and complex environments.

1.2 OBJECTIVES

The main objective of this PhD thesis is to develop a rigorous mobile mapping approach by incorporating a rigorous fisheye lens model to fully exploit the potential of hyperhemispherical fisheye images for real-time 3D navigation and reconstruction. The research focuses on the application and evaluation of this approach, particularly in GNSS-denied and corridor-like environments, with real-time performance.

The mobile mapping method integrates vSLAM techniques and precise photogrammetric models for omnidirectional sensors. A mobile platform equipped with fisheye optical sensors offering a 360° FoV is utilised. The system provides trajectory estimates, which are computed in real time alongside a sparse point cloud using multiprocessing techniques that exploit both CPU and GPU resources. The main contributions of this work address the following two challenges: (1) fisheye lens and omnidirectional dual-camera geometries, and (2) real-time vSLAM performance in outdoor, GNSS-denied, and corridor-like environments without loop closure.

1.3 HYPOTHESES

To address these objectives, we hypothesise that: (H1) the application of rigorous fisheye mathematical models (e.g., equisolid-angle) in a dual omnidirectional camera system with an expanded FoV can enhance the accuracy of vSLAM trajectory solutions compared to generic camera models, such as EUCM, by enabling the tracking of more points and delivering greater consistency in the presence of potential failures; (H2) real-time performance in outdoor, corridor-like environments, where external sensor data is difficult to obtain and revisiting the starting point is challenging, can be improved by utilizing precise mathematical models and embedded GPU boards, particularly in portable mobile mapping platforms (e.g., backpacks and robots), leading to a more robust and consistent vSLAM solution.

To investigate these hypotheses, this work presents an adapted version of ORB-SLAM for fisheye images using the equisolid-angle model. It includes modifications to assess real-time ORB-SLAM performance on Nvidia Jetson TX2 single board computer, following the approach of Aldegheri et al. (2019). The system's performance was evaluated with real data from a backpack-mounted system equipped with a Ricoh Theta S omnidirectional camera, which captures a 360º view of the scene with hyperhemispherical geometry (Campos et al., 2018b).

1.4 THESIS STRUCTURE

This thesis is divided into six chapters: Chapter 1 introduces the context of this thesis, its objectives, and hypotheses. Chapter 2 provides the theoretical background for the key topics addressed in the subsequent chapters. It covers the terms and approaches used in Computer Vision, Photogrammetry, and Robotics, three of the main fields in which vSLAM is developed and applied, thereby contextualising the methodologies proposed in this thesis. Because vSLAM and its constituent techniques are studied across these fields, both the overall framework and its underlying methods are often referred to by different names, or share the same name while carrying different meanings, depending on the area of application. To ensure consistency, the definitions introduced in this chapter are adopted as the common reference throughout the remainder of this thesis.

Chapters 3 to 5 present the main findings of this thesis. These chapters provide a synthesis of the research reported in the derived publications of this thesis by Garcia et al. (2021), Garcia et al. (2023), Garcia et al. (2024a), Garcia et al. (2024b), and Garcia et al. (2025).

Chapter 3 consolidates two complementary studies on adapting ORB-SLAM for fisheye cameras in a single-sensor configuration. Experiment I (Garcia et al., 2023) provides a comparative analysis of a generic fisheye model and the rigorous equisolid-angle model within the ORB-SLAM framework. Experiment II (Garcia et al., 2024b) introduces a real-time vSLAM approach using the rigorous equisolid-angle projection model, with performance assessed in real-time on the Nvidia Jetson TX2 board over a 140-meter outdoor trajectory. Both studies identified that keypoints in the sky region are a major source of trajectory drift in outdoor scenes, motivating the incorporation of a semantic filtering stage.

Chapter 4 presents a machine learning approach for feature-based filtering of sky keypoints in ORB-SLAM. Building on a preliminary study by Garcia et al. (2021) that showed CNN-based semantic segmentation is computationally expensive for real-time deployment on lightweight platforms, the proposed Machine Learning Fisheye Sky Filter (ML-FSF) classifies already-extracted keypoints as sky or non-sky using a model trained on the WoodScape dataset. Results indicate a potential improvement of up to 40% in trajectory estimation accuracy with only a 20% increase in computation time (Garcia et al., 2024a; Garcia et al., 2025).

Chapter 5 presents the dual equisolid ORB-SLAM, a real-time vSLAM solution for a dual omnidirectional camera system implemented on the Nvidia Jetson TX2 board. Experiments compare the dual equisolid ORB-SLAM against the single-sensor approach from Chapter 3, using 360-degree image frames from the Ricoh Theta S camera in a mobile backpack platform.

The summary of results, final remarks, conclusions, and future recommendations are presented in Chapter 6.

CHAPTER 6
FINAL REMARKS

In this thesis, we initially hypothesised that the application of rigorous fisheye mathematical models in a dual omnidirectional camera system with an expanded FoV could enhance the accuracy of vSLAM trajectory solutions compared to generic camera models, thereby providing greater consistency in the presence of potential failures. The first hypothesis was initially addressed in Chapters 3 and 4, where we identified that the rigorous photogrammetric mathematical modelling, which accounts for the image formation process applied to ORB-SLAM, proved superior performance to the simplified and generic models. Using only one sensor, we confirmed the first hypothesis, as the equisolid ORB-SLAM solution was able to track more points and deliver more consistent results compared to the generic model. However, we faced challenges in the outdoor scenario, where a significant portion of the detected keypoints negatively affected the solution due to their excessive distance from the camera, often leading to instability and lost tracking. To address this, in Chapter 4, we focused on integrating a semantic layer into our proposed solution to filter out sky points (ML-FSF), which were degrading the sequential estimation process. This enhancement enabled us to support the first hypothesis, achieving a more robust and stable positioning solution.

The first hypothesis was fully addressed in Chapter 5 by extending the solution from Chapter 4 to integrate both sensors of the Ricoh Theta S camera into the dual equisolid ORB-SLAM. This enhancement aimed to make the proposed approach more robust in scenarios where distinguishable points are not always present in a single sensor. With a 360º view of the scene, the likelihood of detecting more keypoints is significantly higher, enabling a more resilient solution to challenges and potential failures. However, new challenges emerged when using both sensors simultaneously for real-time positioning estimation, mainly the definition of a scale to apply to the cameras' offsets.

Our second hypothesis states that real-time performance in outdoor, corridor-like environments, where external sensor data is difficult to obtain and revisiting the starting point is challenging, can be improved by utilising precise mathematical models and embedded GPU boards, particularly in portable mobile mapping platforms (e.g., backpacks and robots), leading to a more robust and consistent vSLAM solution. This hypothesis has been addressed from the beginning of the methodology in Chapters 3, 4, and 5. Since Chapter 3, we have shown that implementing the proposed methodology with a rigorous mathematical model on a dedicated GPU-enabled board (Jetson TX2) has resulted in a frame processing time that enables real-time vSLAM applications on mobile platforms. Nevertheless, by integrating all the proposed approaches developed for processing images from a dual omnidirectional camera system with an expanded FoV in Chapter 5, we faced some challenges that prevented us from fully satisfying the second hypothesis. The processing time per frame exceeded the threshold required for real-time implementation, highlighting the need for optimizations to reduce computational cost, particularly by minimizing the number of keypoints necessary for effective tracking in the solution.

Although the use of two sensors (Ricoh Theta S) did not fully achieve the initial objective due to previously mentioned challenges, the approach's potential can be recognized, along with opportunities to address the faced issues and fully achieve the intended goals. The following considerations highlight potential improvements for enhancing and advancing the proposed approach:

(1) We utilised frames extracted from video captured by the Ricoh Theta S. The video has a lower resolution compared to static images, and resolution directly impacts feature-based vSLAM, as higher-quality images enable the extraction of distinguishable points more efficiently. The proposed approach could be enhanced by employing a higher-resolution camera, such as the Ricoh Theta Z, which offers superior resolution and quality compared to the Ricoh Theta S, or by exploring other omnidirectional cameras currently available.

(2) When processing images from both sensors simultaneously, it was not possible to maintain a 5 fps capture rate, as the average processing time per frame nearly doubled due to the increased number of keypoints processed in each frame. However, it is feasible to reduce the capture rate, as the higher number of keypoints contributes to greater stability between frames, thus preventing lost tracking caused by a lack of trackable keypoints in consecutive frames. Another aspect that can be explored is the correction of distortion parameters in the adopted model. Currently, distortions are corrected by calculating it for each point, but this can be optimized and accelerated by using a lookup table (LUT) with precomputed distortion estimates for each pixel (Peng, Sun and Wang, 2023).

(3) Continuing with the theme of computational optimisation to improve response time, another refinement that can be implemented involves the mathematical model. In this work, we used a single EOP for each frame to reduce the number of parameters to be estimated, thereby lowering computational costs. Additionally, considering the known geometry of the camera used in the experiments, it is worth exploring and developing

alternatives to employ only one single set of IOPs as well, by adjusting both the mathematical model and the calibration process.

(4) The main challenge we encountered in identifying issues and necessary adjustments in the methodology was that ORB-SLAM assigns a local reference frame to the trajectory, determined by randomly selected initialisation points. This complicates comparisons with the reference adopted in UTM coordinates (E, N) and ellipsoidal height (h). Bringing the solution to the scale and reference coordinate system of the real world is a critical consideration and can be approached in various ways, such as incorporating targets with known dimensions along the trajectory or using baseline constraints. Another notable observation during the experiments was that as drift accumulates in the trajectory, there is deformation and scale alteration when loop closure is absent, as previously discussed by Yang et al. (2024). It is also possible to integrate data from additional sensors, such as GNSS receivers, when such measurements can be obtained based on the scenario requirements.

(5) Given the arbitrary definition of the real-world scale, it is advisable to explore the mathematical modelling in the dual approach to address the challenges encountered along the trajectory. These include significant variations in scene scale, the impact of accumulated drift on reference scale consistency, and the calibration and estimation of ROPs. One aspect to consider is that, although the images from both sensors of the Ricoh Theta S camera are treated as a single frame in the mathematical modelling to optimize nonlinear estimation calculations, it would be more appropriate to handle the 2D keypoints individually for each image. This approach would allow for the application of search constraints and keypoint tracking strategies that best account for both sensors.

(6) Mur-Artal and Tardos (2017) mentioned that the loop closing thread is designed to reduce drift accumulated during traversing when returning to a previously mapped area. The place recognition module matches a recent keyframe with an earlier one, and this match is validated by computing a rigid body transformation to align the matched points between the keyframes. Finally, a global optimization process is performed to minimize the trajectory error. Essentially, points with previously known coordinates are identified to apply global optimization. As mentioned earlier, our objective is to explore and simulate corridor-like scenarios, where revisiting a location not feasible. In such cases, it is possible to use targets with known coordinates in the environment to perform loop closure, particularly at the middle and end of the trajectory. This would contribute to a more consistent trajectory with reduced discrepancies.

(7) Outdoor vSLAM faces numerous uncontrolled challenges that can negatively impact the solution, such as huge depth variations with distant objects in the scene and dynamic elements. To advance toward increasingly robust methods, it is essential to integrate intelligence into data capture through machine learning techniques. This approach can enhance the solution's resilience and robustness against uncontrolled factors. We confirmed this by implementing the ML-FSF filter, which significantly enhanced the stability of the solution, reduced instances of lost tracking, and consequently improved accuracy.

(8) It would be appropriate to test the approach presented in Chapter 5 in different scenarios, such as environments with less variation in scene scale, to analyse the behaviour of the solution. Additionally, we could further investigate the mathematical calculations to identify the specific factor that caused the solution to degrade significantly after reaching just over 40 m along the path of the experiment in Chapter 5. The adjustment of the frame rate is also an aspect to be explored to determine whether an increased number of frames per second would improve the estimation behaviour, particularly by better handling situations with high variations in scene scale and reducing trajectory drift.

REFERENCES

ABASPUR KAZEROUNI, I.; FITZGERALD, L.; DOOLY, G.; TOAL, D. A survey of state-of-the-art on visual SLAM. **Expert Systems with Applications**, v. 205, p. 117734, 2022. https://doi.org/10.1016/j.eswa.2022.117734.

ABDEL-AZIZ, Y. I.; KARARA, H. M. Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-Range Photogrammetry. **Photogrammetric Engineering & Remote Sensing**, v. 81, n. 2, p. 103–107, 2015. https://doi.org/10.14358/PERS.81.2.103.

ABRAHAM, S.; FÖRSTNER, W. Fish-eye-stereo calibration and epipolar rectification. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 59, n. 5, p. 278–288, 2005. https://doi.org/10.1016/j.isprsjprs.2005.03.001.

AHMED, M. F.; MASOOD, K.; FREMONT, V.; FANTONI, I. Active SLAM: A Review on Last Decade. **Sensors**, v. 23, n. 19, p. 8097, 2023. https://doi.org/10.3390/s23198097.

ALDEGHERI, S.; BOMBIERI, N.; BLOISI, D. D.; FARINELLI, A. Data Flow ORB-SLAM for Real-time Performance on Embedded GPU Boards. 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). **Proceedings**... In: 2019 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS). Macau, China: IEEE, 2019. https://doi.org/10.1109/IROS40897.2019.8967814.

ALSADIK, B.; KARAM, S. The Simultaneous Localization and Mapping (SLAM)-An Overview. **Journal of Applied Science and Technology Trends**, v. 2, n. 04, p. 120–131, 2021. https://doi.org/10.38094/jastt204117.

BARROS, A. M.; MICHEL, M.; MOLINE, Y.; CORRE, G.; CARREL, F. A Comprehensive Survey of Visual SLAM Algorithms. **Robotics**, v. 11, n. 1, p. 24, 2022. https://doi.org/10.3390/robotics11010024.

BAYOUD, F. A.; SKALOUD, J.; BERTRAND, M. Photogrammetry-derived Navigation Parameters for INS Kalman Filter Updates. XXth ISPRS Congress. **Proceedings**... In: XXTH ISPRS CONGRESS. Istanbul (Turkey): 2004.

BEGHDADI, A.; MALLEM, M. A comprehensive overview of dynamic visual SLAM and deep learning: concepts, methods and challenges. **Machine Vision and Applications**, v. 33, art. 54, 2022. https://doi.org/10.1007/s00138-022-01306-w.

BREIMAN, L. Random Forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001. https://doi.org/10.1023/A:1010933404324.

BRESSON, G.; ALSAYED, Z.; YU, L.; GLASER, S. Simultaneous Localization and Mapping: A Survey of Current Trends in Autonomous Driving. **IEEE Transactions on Intelligent Vehicles**, v. 2, n. 3, p. 194–220, 2017. https://doi.org/10.1109/TIV.2017.2749181.

BROWN, D. C. Close-range camera calibration. **Photogrammetric Engineering**, v. 37, n. 8, p. 855–866, 1971.

CADENA, C.; CARLONE, L.; CARRILLO, H.; LATIF, Y.; SCARAMUZZA, D.; NEIRA, J.; REID, I.; LEONARD, J. J. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. **IEEE Transactions on Robotics**, v. 32, n. 6, p. 1309–1332, 2016.

CAMARGO, P. O. Ajustamento de Observações. Course notes. Presidente Prudente: UNESP, 2000.

CAMPOS, C.; ELVIRA, R.; GÓMEZ RODRÍGUEZ, J. J.; MONTIEL, J. M. M.; TARDÓS, J. D. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM. **arXiv**, 2007.11898, 2020. https://doi.org/10.48550/ARXIV.2007.11898.

CAMPOS, M. B. **A Portable Mobile Terrestrial System With Omnidirectional Camera For Close Range Applications**. PhD Thesis—Presidente Prudente: Faculdade de Ciências e Tecnologia de Presidente Prudente, 2019.

CAMPOS, M. B.; TOMMASELLI, A. M. G.; MARCATO JUNIOR, J.; HONKAVAARA, E. Geometric model and assessment of a dual-fisheye imaging system. **The Photogrammetric Record**, v. 33, n. 162, p. 243–263, 2018a. https://doi.org/10.1111/phor.12240.

CAMPOS, M.; TOMMASELLI, A.; HONKAVAARA, E.; PROL, F.; KAARTINEN, H.; EL ISSAOUI, A.; HAKALA, T. A Backpack-Mounted Omnidirectional Camera with Off-the-Shelf Navigation Sensors for Mobile Terrestrial Mapping: Development and Forest Application. **Sensors**, v. 18, n. 3, p. 827, 2018b. https://doi.org/10.3390/s18030827.

CARUSO, D.; ENGEL, J.; CREMERS, D. Large-scale direct SLAM for omnidirectional cameras. 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). **Proceedings**... In: 2015 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS). Hamburg, Germany: IEEE, 2015. https://doi.org/10.1109/IROS.2015.7353366.

CASTANHEIRO, L. F. **Geometric Model Of A Dual-Fisheye System Composed Of Hyper-Hemispherical Lenses**. Master Thesis—Presidente Prudente: Faculdade de Ciências e Tecnologia de Presidente Prudente, 2020.

CASTANHEIRO, L. F.; TOMMASELLI, A. M. G.; BERVEGLIERI, A.; CAMPOS, M. B.; JUNIOR, J. M. Modeling Hyperhemispherical Points and Calibrating a Dual-Fish-Eye System for Close-Range Applications. **Photogrammetric Engineering & Remote Sensing**, v. 87, n. 5, p. 375–384, 2021. https://doi.org/10.14358/PERS.87.5.375.

CHAVES, S. M.; KIM, A.; GALCERAN, E.; EUSTICE, R. M. Opportunistic Sampling-based Active Visual SLAM for Underwater Inspection. **Autonomous Robots**, v. 40, n. 7, p. 1245–1265, 2016. https://doi.org/10.1007/s10514-016-9597-6.

CHEBROLU, N.; LOTTES, P.; LÄBE, T.; STACHNISS, C. Robot Localization Based on Aerial Images for Precision Agriculture Tasks in Crop Fields. 2019 International Conference on Robotics and Automation (ICRA). **Proceedings**... In: 2019 INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA). Montreal, Canada: IEEE, 2019. p. 1787–1793. https://doi.org/10.1109/ICRA.2019.8794030.

CHEN, L.; PAPANDREOU, G.; KOKKINOS, I.; MURPHY, K.; YUILLE, A. L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. **arXiv**, 1606.00915, 2016. https://doi.org/10.48550/ARXIV.1606.00915.

CHEN, W.; SHANG, G.; JI, A.; ZHOU, C.; WANG, X.; XU, C.; LI, Z.; HU, K. An Overview on Visual SLAM: From Tradition to Semantic. **Remote Sensing**, v. 14, n. 13, p. 3010, 2022. https://doi.org/10.3390/rs14133010.

CHEN, X.; WEIDONG, H.; LEFENG, Z.; SHI, Z.; LI, M. Integration of Low-Cost GNSS and Monocular Cameras for Simultaneous Localization and Mapping. **Sensors**, v. 18, n. 7, p. 2193–2211, 2018.

CIVERA, J.; DAVISON, A. J.; MONTIEL, J. M. M. Inverse Depth Parametrization for Monocular SLAM. **IEEE Transactions on Robotics, v. 24, n. 5, p. 932–945, 2008.** https://doi.org/10.1109/TRO.2008.2003276.

CONCHA, E. K. T.; PITTOL, D.; WESTHAUSER, R.; KOLBERG, M.; MAFFEI, R.; PRESTES, E. Map Point Optimization in Keyframe-Based SLAM using Covisibility Graph and Information Fusion. 2019 19th International Conference on Advanced Robotics (ICAR). **Proceedings**...Belo Horizonte, Brazil: IEEE, 2019. https://doi.org/10.1109/ICAR46387.2019.8981653.

CONRADY, A. E. Decentred Lens-Systems. **Monthly Notices of the Royal Astronomical Society**, v. 79, n. 5, p. 384–390, 1919. https://doi.org/10.1093/mnras/79.5.384.

COURBON, J.; MEZOUAR, Y.; ECKT, L.; MARTINET, P. A generic fisheye camera model for robotic applications. 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems. **Proceedings**... In: 2007 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS. San Diego, CA: IEEE, 2007. https://doi.org/10.1109/IROS.2007.4399233.

DALMOLIN, Q. **Ajustamento por Mínimos Quadrados**. 1st ed. Curitiba: Imprensa Universitária – UFPR, 2002.

DAVISON, A. J. Real-time simultaneous localisation and mapping with a single camera. Proceedings Ninth IEEE International Conference on Computer Vision. **Proceedings**...Nice, France: IEEE, 2003. https://doi.org/10.1109/ICCV.2003.1238654.

DAVISON, A. J.; REID, I. D.; MOLTON, N. D.; STASSE, O. MonoSLAM: Real-Time Single Camera SLAM. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 29, n. 6, p. 1052–1067, 2007. https://doi.org/10.1109/TPAMI.2007.1049.

DISSANAYAKE, M. W. M. G.; NEWMAN, P.; CLARK, S.; DURRANT-WHYTE, H. F.; CSORBA, M. A solution to the simultaneous localization and map building (SLAM) problem. **IEEE Transactions on Robotics and Automation**, v. 17, n. 3, p. 229–241, 2001. https://doi.org/10.1109/70.938381.

DURRANT-WHYTE, H.; BAILEY, T. Simultaneous localization and mapping: part I. **IEEE robotics & automation magazine**, v. 13, n. 2, p. 99–110, 2006.

ENGEL, J.; SCHÕPS, T.; CREMERS, D. LSD-SLAM: Large-Scale Direct Monocular SLAM. In: FLEET, D.; PAJDLA, T.; SCHIELE, B.; TUYTELAARS, T. (Eds.). **Computer Vision – ECCV 2014**. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2014. v. 8690p. 834–849. https://doi.org/10.1007/978-3-319-10605-2_54.

FAUGERAS, O. D.; LUSTMAN, F. Motion and structure from motion in a piecewise planar environment. **International Journal of Pattern Recognition and Artificial Intelligence**, v. 02, n. 03, p. 485–508, 1988. https://doi.org/10.1142/S0218001488000285.

FISCHLER, M. A.; BOLLES, R. C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. **Communications of the ACM**, v. 24, n. 6, p. 381–395, 1981. https://doi.org/10.1145/358669.358692.

FORSTER, C.; PIZZOLI, M.; SCARAMUZZA, D. SVO: Fast semi-direct monocular visual odometry. 2014 IEEE International Conference on Robotics and Automation (ICRA). **Proceedings**... In: 2014 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA). Hong Kong, China: IEEE, 2014. https://doi.org/10.1109/ICRA.2014.6906584.

FORSTNER, W. A feature based correspondence algorithm for image matching. **Int. Arch. Photogrammetry**, v. 26, n. 3, p. 150–166, 1986.

FORSTNER, W.; WROBEL, B. P. **Photogrammetric Computer Vision: Statistics, Geometry, Orientation and Reconstruction**. [s.l.] Springer, 2016.

FRASER, C. S. Digital camera self-calibration. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 52, n. 4, p. 149–159, 1997. https://doi.org/10.1016/S0924-2716(97)00005-1.

FUENTES-PACHECO, J.; RUIZ-ASCENCIO, J.; RENDÓN-MANCHA, J. M. Visual simultaneous localization and mapping: a survey. **Artificial Intelligence Review**, v. 43, n. 1, p. 55–81, 2015. https://doi.org/10.1007/s10462-012-9365-8.

GALO, M.; TOMMASELLI, A. M. G. Calibração de Câmaras. Unpublished, 2011. Available at: <https://www.researchgate.net/doi/10.13140/2.1.2759.7764>. Accessed: 9 Jan. 2026.

GALVEZ-LÓPEZ, D.; TARDOS, J. D. Bags of Binary Words for Fast Place Recognition in Image Sequences. **IEEE Transactions on Robotics**, v. 28, n. 5, p. 1188–1197, 2012. https://doi.org/10.1109/TRO.2012.2197158.

GAO, X.; ZHANG, T. **Introduction to Visual SLAM: From Theory to Practice**. Singapore: Springer Singapore, 2021. https://doi.org/10.1007/978-981-16-4939-4.

GARCIA, T. A. C.; CAMPOS, M. B.; CASTANHEIRO, L. F.; TOMMASELLI, A. M. G. A Proposal to Integrate ORB-SLAM Fisheye and Convolutional Neural Networks for Outdoor Terrestrial Mobile Mapping. 2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). **Proceedings**... In: 2021 IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM (IGARSS). Brussels, Belgium: IEEE, 2021. p. 578–581. https://doi.org/10.1109/IGARSS47720.2021.9554752.

GARCIA, T. A. C.; TOMMASELLI, A. M. G.; CASTANHEIRO, L. F.; CAMPOS, M. B. Comparison of Two Mathematical Models for Fisheye Cameras Applied In the ORB-SLAM. THE INTERNATIONAL SYMPOSIUM ON MOBILE MAPPING TECHNOLOGY (MMT) 2023. **Proceedings**... In: THE INTERNATIONAL SYMPOSIUM ON MOBILE MAPPING TECHNOLOGY (MMT) 2023. Padua (Italy): 2023.

GARCIA, T. A. C.; TOMMASELLI, A. M. G.; CASTANHEIRO, L. F.; CAMPOS, M. B. Filtering Keypoints with ORB Features. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. **Proceedings**...Belém (Brazil): 2024a. v. XLVIII-3-2024, p. 177–182. https://doi.org/10.5194/isprs-archives-XLVIII-3-2024-177-2024.

GARCIA, T. A. C.; TOMMASELLI, A. M. G.; CASTANHEIRO, L. F.; CAMPOS, M. B. A photogrammetric approach for real-time visual SLAM applied to an omnidirectional system. **The Photogrammetric Record**, p. phor.12494, 2024b. https://doi.org/10.1111/phor.12494.

GARCIA, T. A. C.; TOMMASELLI, A. M. G.; CASTANHEIRO, L. F.; CAMPOS, M. B. A Machine Learning Approach to Filtering ORB Features in Outdoor Real-time Fisheye vSLAM. **To be published**, 2025.

GARFORTH, J.; WEBB, B. Visual Appearance Analysis of Forest Scenes for Monocular SLAM. 2019 International Conference on Robotics and Automation (ICRA). **Proceedings**... In: 2019 INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA). Montreal, Canada: IEEE, 2019. p. 1794–1800. https://doi.org/10.1109/ICRA.2019.8793771.

GÉRON, A. **Hands-On Machine Learning with Scikit-Learn, Keras, and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems**. 2rd. ed. Canada: O'Reilly Media, 2019.

GEYER, C.; DANIILIDIS, K. A Unifying Theory for Central Panoramic Systems and Practical Applications. Proceedings of the 6th European Conference on Computer Vision-Part II. **Proceedings**...: ECCV '00. In: 6TH EUROPEAN CONFERENCE ON COMPUTER VISION-PART II. Berlin, Heidelberg: Springer-Verlag, 2000.

GOLUB, G. H.; VAN LOAN, C. F. **Matrix Computations**. 2nd ed. [s.l.] The Johns Hopkins University Press, 1989.

GRUEN, A. Towards real-time photogrammetry. **Photogrammetria**, v. 42, n. 5–6, p. 209–244, 1988. https://doi.org/10.1016/0031-8663(88)90003-8.

GRUSSENMEYER, P.; KHALIL, O. A. SOLUTIONS FOR EXTERIOR ORIENTATION IN PHOTOGRAMMETRY: A REVIEW. **The Photogrammetric Record**, v. 17, n. 100, p. 615–634, 2002. https://doi.org/10.1111/j.1477-9730.2002.tb01907.x.

HABIB, A.; MORGAN, M. Stability Analysis and Geometric Calibration of Off-the-Shelf Digital Cameras. **Photogrammetric Engineering & Remote Sensing**, v. 71, n. 6, p. 733–741, 2005. https://doi.org/10.14358/PERS.71.6.733.

HAGGREN, H. Real-Time Photogrammetry as used for machine vision applications. **Real-Time Photogrammetry - A New Challenge**, v. 26, n. 5, p. 374–382, 1986.

HARTLEY, R.; ZISSERMAN, A. **Multiple View Geometry in Computer Vision**. 2nd ed. [s.l.] Cambridge University Press, 2004. https://doi.org/10.1017/CBO9780511811685.

HUGHES, C.; DENNY, P.; JONES, E.; GLAVIN, M. Accuracy of fish-eye lens models. **Applied Optics**, v. 49, n. 17, p. 3338, 2010. https://doi.org/10.1364/AO.49.003338.

KANNALA, J.; BRANDT, S. S. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 28, n. 8, p. 1335–1340, 2006. https://doi.org/10.1109/TPAMI.2006.153.

KATO, S.; KITAMURA, M.; SUZUKI, T.; AMANO, Y. NLOS Satellite Detection Using a Fish-Eye Camera for Improving GNSS Positioning Accuracy in Urban Area. **Journal of Robotics and Mechatronics**, v. 28, n. 1, p. 31–39, 2016. https://doi.org/10.20965/jrm.2016.p0031.

KHOMUTENKO, B.; GARCIA, G.; MARTINET, P. An Enhanced Unified Camera Model. **IEEE Robotics and Automation Letters**, v. 1, n. 1, p. 137–144, 2016. https://doi.org/10.1109/LRA.2015.2502921.

KLEIN, G.; MURRAY, D. Parallel Tracking and Mapping for Small AR Workspaces. 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality. **Proceedings**... In: 2007 6TH IEEE INTERNATIONAL SYMPOSIUM ON MIXED AND AUGMENTED REALITY (ISMAR). Nara, Japan: IEEE, 2007. https://doi.org/10.1109/ISMAR.2007.4538852.

KUMAR, V. R.; EISING, C.; WITT, C.; YOGAMANI, S. Surround-view Fisheye Camera Perception for Automated Driving: Overview, Survey and Challenges. **arXiv**, 2205.13281, 2022. https://arxiv.org/abs/2205.13281.

KUMMERLE, R.; GRISETTI, G.; STRASDAT, H.; KONOLIGE, K.; BURGARD, W. G2o: A general framework for graph optimization. 2011 IEEE International Conference on Robotics and Automation. **Proceedings**... In: 2011 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA). Shanghai, China: IEEE, 2011. https://doi.org/10.1109/ICRA.2011.5979949.

LEUTENEGGER, S.; LYNEN, S.; BOSSE, M.; SIEGWART, R.; FURGALE, P. Keyframe-based visual–inertial odometry using nonlinear optimization. **The International Journal of Robotics Research**, v. 34, n. 3, p. 314–334, 2015. https://doi.org/10.1177/0278364914554813.

LEVENBERG, K. A method for the solution of certain non-linear problems in least squares. **Quarterly of Applied Mathematics**, v. 2, n. 2, p. 164–168, 1944. https://doi.org/10.1090/qam/10666.

LI, A.; RUAN, X.; HUANG, J.; ZHU, X.; WANG, F. Review of vision-based Simultaneous Localization and Mapping. 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). **Proceedings**...Chengdu, China: IEEE, 2019. https://doi.org/10.1109/ITNEC.2019.8729285.

LI, J.; WANG, X.; LI, S. Spherical-Model-Based SLAM on Full-View Images for Indoor Environments. **Applied Sciences**, v. 8, n. 11, p. 2268, 2018. https://doi.org/10.3390/app8112268.

LI, S.; ZHANG, D.; XIAN, Y.; LI, B.; ZHANG, T.; ZHONG, C. Overview of deep learning application on visual SLAM. **Displays**, v. 74, art. 102298, 2022. https://doi.org/10.1016/j.displa.2022.102298.

LIU, S.; GUO, P.; FENG, L.; YANG, A. Accurate and Robust Monocular SLAM with Omnidirectional Cameras. **Sensors**, v. 19, n. 20, p. 4494, 2019. https://doi.org/10.3390/s19204494.

LONGUET-HIGGINS, H. C. A computer algorithm for reconstructing a scene from two projections. **Nature**, v. 293, n. 5828, p. 133–135, 1981. https://doi.org/10.1038/293133a0.

LOWE, D. G. Distinctive image features from scaleinvariant keypoints. **International Journal of Computer Vision**, v. 60, p. 91–110, 2004.

MARQUARDT, D. W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. **Journal of the Society for Industrial and Applied Mathematics**, v. 11, n. 2, p. 431–441, 1963. https://doi.org/10.1137/0111030.

MATSUKI, H.; VON STUMBERG, L.; USENKO, V.; STUCKLER, J.; CREMERS, D. Omnidirectional DSO: Direct Sparse Odometry with Fisheye Cameras. **arXiv**, 1808.02775, 2018. https://doi.org/10.48550/ARXIV.1808.02775.

MENNA, F.; BATTISTI, R.; NOCERINO, E.; REMONDINO, F. FROG: A Portable Underwater Mobile Mapping System. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. **Proceedings**...Padua (Italy): 2023. v. XLVIII-1/W1-2023, p. 295–302. https://doi.org/10.5194/isprs-archives-XLVIII-1-W1-2023-295-2023.

MIKHAIL, E. M.; BETHEL, J. S.; MCGLONE, J. C. **Introduction to modern photogrammetry**. New York : Chichester: Wiley, 2001.

MONIWA, H. The concept of “photo-variant” self-calibration and its application in block adjustment with bundles. **Photogrammetria**, v. 36, n. 1, p. 11–29, 1981. https://doi.org/10.1016/0031-8663(81)90006-5.

MORAVEC, H. **Obstacle avoidance and navigation in the real world by a seeing robot rover**. PhD Thesis—Stanford, CA: Stanford University, 1980.

MORÉ, J. J. The Levenberg-Marquardt algorithm: Implementation and theory. In: WATSON, G. A. (Ed.). **Numerical Analysis**. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1978. v. 630p. 105–116. https://doi.org/10.1007/BFb0067700.

MUR-ARTAL, R.; MONTIEL, J. M. M.; TARDOS, J. D. ORB-SLAM: A versatile and accurate monocular SLAM system. **IEEE Transactions on Robotics**, v. 31, n. 5, p. 1147–1163, 2015.

MUR-ARTAL, R.; TARDOS, J. D. ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras. **arXiv**, 1610.06475, 2016. https://doi.org/10.48550/ARXIV.1610.06475.

MUR-ARTAL, R.; TARDOS, J. D. Visual-Inertial Monocular SLAM With Map Reuse. **IEEE Robotics and Automation Letters**, v. 2, n. 2, p. 796–803, 2017. https://doi.org/10.1109/LRA.2017.2653359.

NEWCOMBE, R. A.; LOVEGROVE, S. J.; DAVISON, A. J. DTAM: Dense tracking and mapping in real-time. 2011 International Conference on Computer Vision. **Proceedings**... In:

2011 IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV). Barcelona, Spain: IEEE, 2011. https://doi.org/10.1109/ICCV.2011.6126513.

NICE, K. A.; WIJNANDS, J. S.; MIDDEL, A.; WANG, J.; QIU, Y.; ZHAO, N.; THOMPSON, J.; ASCHWANDEN, G. D. P. A.; ZHAO, H.; STEVENSON, M. Sky pixel detection in outdoor imagery using an adaptive algorithm and machine learning. **Urban Climate**, v. 31, p. 100572, 2020. https://doi.org/10.1016/j.uclim.2019.100572.

PENG, H.; SUN, C.; WANG, P. The Design and Implementation of the Lookup Table in Distortion Correction for the Optical System. **IEEE Access**, v. 11, p. 44335–44342, 2023. https://doi.org/10.1109/ACCESS.2023.3272615.

PU, H.; LUO, J.; WANG, G.; HUANG, T.; LIU, H.; LUO, J. Visual SLAM Integration With Semantic Segmentation and Deep Learning: A Review. **IEEE Sensors Journal**, v. 23, n. 19, p. 22119–22138, 2023. https://doi.org/10.1109/JSEN.2023.3306371.

RAY, S. F. **Applied photographic optics: lenses and optical systems for photography, film, video, and electronic imaging**. 2nd ed. Oxford ; Boston: Focal Press, 1994.

RUBLEE, E.; RABAUD, V.; KONOLIGE, K.; BRADSKI, G. ORB: An efficient alternative to SIFT or SURF. IEEE International Conference on Computer Vision. **Proceedings**...2011.

SCARAMUZZA, D.; FRAUNDORFER, F. Visual Odometry [Tutorial]. **IEEE Robotics & Automation Magazine**, v. 18, n. 4, p. 80–92, 2011. https://doi.org/10.1109/MRA.2011.943233.

SCARAMUZZA, D.; FRAUNDORFER, F.; SIEGWART, R. Real-time monocular visual odometry for on-road vehicles with 1-point RANSAC. 2009 IEEE International Conference on Robotics and Automation. **Proceedings**... In: 2009 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA). Kobe: IEEE, 2009. https://doi.org/10.1109/ROBOT.2009.5152255.

SCHNEIDER, D.; SCHWALBE, E.; MAAS, H.-G. Validation of geometric models for fisheye lenses. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 64, n. 3, p. 259–266, 2009. https://doi.org/10.1016/j.isprsjprs.2009.01.001.

SCHNEIDER, J.; FORSTNER, W. Bundle Adjustment and System Calibration with Points at Infinity for Omnidirectional Camera Systems. **Photogrammetrie - Fernerkundung - Geoinformation**, v. 2013, n. 4, p. 309–321, 2013. https://doi.org/10.1127/1432-8364/2013/0179.

SHAO, C.; ZHANG, C.; FANG, Z.; YANG, G. A Deep Learning-Based Semantic Filter for RANSAC-Based Fundamental Matrix Calculation and the ORB-SLAM System. **IEEE Access**, v. 8, p. 3212–3223, 2020. https://doi.org/10.1109/ACCESS.2019.2962268.

SHEN, Y.; WANG, Q. Sky Region Detection in a Single Image for Autonomous Ground Robot Navigation. **International Journal of Advanced Robotic Systems**, v. 10, n. 10, p. 362, 2013. https://doi.org/10.5772/56884.

SHOBHA, G.; RANGASWAMY, S. Machine Learning. In: **Handbook of Statistics**. [s.l.] Elsevier, 2018. v. 38p. 197–228. https://doi.org/10.1016/bs.host.2018.07.004.

SIBLEY, G.; MEI, C.; REID, I.; NEWMAN, P. Adaptive Relative Bundle Adjustment. In: TRINKLE, J.; MATSUOKA, Y.; CASTELLANOS, J. A. (Eds.). **Robotics**. [s.l.] The MIT Press, 2010. p. 177–184. https://doi.org/10.7551/mitpress/8727.003.0024.

SONG, H.; LIU, C.; DAI, H. BundledSLAM: An Accurate Visual SLAM System Using Multiple Cameras. **arXiv**, 2403.19886, 2024. Available at: <https://arxiv.org/abs/2403.19886>. Accessed: 10 Oct. 2024.

STRASDAT, H.; DAVISON, A. J.; MONTIEL, J. M. M.; KONOLIGE, K. Double window optimisation for constant time visual SLAM. 2011 International Conference on Computer Vision. **Proceedings**...Barcelona: IEEE, 2011. https://doi.org/10.1109/ICCV.2011.6126517.

STURM, J.; ENGELHARD, N.; ENDRES, F.; BURGARD, W.; CREMERS, D. A Benchmark for the Evaluation of RGB-D SLAM Systems. 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). **Proceedings**... In: 2012 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS). Vilamoura, Portugal: IEEE, 2012. p. 573–580. https://doi.org/10.1109/IROS.2012.6385773.

TAKETOMI, T.; UCHIYAMA, H.; IKEDA, S. Visual SLAM algorithms: a survey from 2010 to 2016. **IPSJ Transactions on Computer Vision and Applications**, v. 9, n. 1, p. 16, 2017. https://doi.org/10.1186/s41074-017-0027-2.

TAO, W.; DAI, Z.; ZHU, X.; LI, F. A sky region segmentation method for outdoor visual-inertial SLAM. 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP). **Proceedings**... In: 2021 6TH INTERNATIONAL CONFERENCE ON INTELLIGENT COMPUTING AND SIGNAL PROCESSING (ICSP). Xi'an, China: IEEE, 2021. https://doi.org/10.1109/ICSP51882.2021.9408771.

TAREEN, S. A. K.; SALEEM, Z. A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK. 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET). **Proceedings**... In: 2018 INTERNATIONAL CONFERENCE ON COMPUTING, MATHEMATICS AND ENGINEERING TECHNOLOGIES (ICOMET). Sukkur: IEEE, 2018. https://doi.org/10.1109/ICOMET.2018.8346440.

TOMMASELLI, A.; GALO, M.; DE MORAES, M.; MARCATO, J.; CALDEIRA, C.; LOPES, R. Generating Virtual Images from Oblique Frames. **Remote Sensing**, v. 5, n. 4, p. 1875–1893, 2013. https://doi.org/10.3390/rs5041875.

TOMMASELLI, A. M. G. **Um Método Recursivo Aplicado ao Problema de Localização em Visão de Máquina**. PhD Thesis—Campinas: Faculdade de Engenharia Elétrica, Universidade Estadual de Campinas, 1993.

TOMMASELLI, A. M. G. Ressecção Espacial e Intersecção Fotogramétrica. Lecture notes. Presidente Prudente: UNESP, 2020.

TOMMASELLI, A. M. G.; GARCIA, T. A. C.; CASTANHEIRO, L. F.; CAMPOS, M. B.; SANTOS, G. H. Effects of Hyper Hemispherical Field in Bundle Adjustment with Fisheye Images. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. XLVIII-1/W1-2023, p. 503–509, 2023. https://doi.org/10.5194/isprs-archives-XLVIII-1-W1-2023-503-2023.

TORRESANI, A.; REMONDINO, F. Videogrammetry VS Photogrammetry for heritage 3D reconstruction. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. XLII-2/W15, p. 1157–1162, 2019. https://doi.org/10.5194/isprs-archives-XLII-2-W15-1157-2019.

TSINTOTAS, K. A.; BAMPIS, L.; GASTERATOS, A. The Revisiting Problem in Simultaneous Localization and Mapping. In: TSINTOTAS, K. A.; BAMPIS, L.; GASTERATOS, A. (Eds.). **Online Appearance-Based Place Recognition and Mapping**. Springer Tracts in Advanced Robotics. Cham: Springer International Publishing, 2022. v. 133p. 1–33. https://doi.org/10.1007/978-3-031-09396-8_1.

UMEYAMA, S. Least-Squares Estimation of Transformation Parameters Between Two Point Patterns. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 13, n. 4, p. 376–380, 1991. https://doi.org/10.1109/34.88573.

USENKO, V.; DEMMEL, N.; CREMERS, D. The Double Sphere Camera Model. **arXiv**, 1807.08957, 2018. https://doi.org/10.48550/ARXIV.1807.08957.

WANG, S.; YUE, J.; DONG, Y.; SHEN, R.; ZHANG, X. Real-time Omnidirectional Visual SLAM with Semi-Dense Mapping. 2018 IEEE Intelligent Vehicles Symposium (IV). **Proceedings**... In: 2018 IEEE INTELLIGENT VEHICLES SYMPOSIUM (IV). Changshu: IEEE, 2018. https://doi.org/10.1109/IVS.2018.8500460.

WANG, Y.; BU, H.; ZHANG, X.; CHENG, J. YPD-SLAM: A Real-Time VSLAM System for Handling Dynamic Indoor Environments. **Sensors**, v. 22, n. 21, art. 8561, 2022. https://doi.org/10.3390/s22218561.

WANG, Y.; DUAN, X.; SUN, Y.; WANG, J. A Visual SLAM Algorithm Based on Image Semantic Segmentation in Dynamic Environment. 2021 4th International Conference on Pattern Recognition and Artificial Intelligence (PRAI). **Proceedings**... In: 2021 4TH INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION AND ARTIFICIAL INTELLIGENCE (PRAI). Yibin, China: IEEE, 2021. https://doi.org/10.1109/PRAI53619.2021.9550800.

WANG, Z.; SHEN, Y.; CAI, B.; SALEEM, M. T. A Brief Review on Loop Closure Detection with 3D Point Cloud. 2019 IEEE International Conference on Real-time Computing and Robotics (RCAR). **Proceedings**... In: 2019 IEEE INTERNATIONAL CONFERENCE ON REAL-TIME COMPUTING AND ROBOTICS (RCAR). Irkutsk, Russia: IEEE, 2019. https://doi.org/10.1109/RCAR47638.2019.9044021.

WELLS, L. A.; CHUNG, W. Vision-Aided Localization and Mapping in Forested Environments Using Stereo Images. **Sensors**, v. 23, n. 16, art. 7043, 2023. https://doi.org/10.3390/s23167043.

WOLF, P. R.; DEWITT, B. A. **Elements of Photogrammetry Book With Application in GIS**. 3. ed. United States: Mcgraw-Hill Education-Europe, 2000.

WU, C. Implementation of Scale Invariant Feature Transform (SIFT). 2021. Available at: <https://github.com/pitzer/SiftGPU>. Accessed: 17 Jan. 2021.

XIAO, L.; WANG, J.; QIU, X.; RONG, Z.; ZOU, X. Dynamic-SLAM: Semantic monocular visual localization and mapping based on deep learning in dynamic environment. **Robotics**

and Autonomous Systems, v. 117, p. 1–16, 2019. https://doi.org/10.1016/j.robot.2019.03.012.

XIE, E.; WANG, W.; YU, Z.; ANANDKUMAR, A.; ALVAREZ, J. M.; LUO, P. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. **arXiv**, 2105.15203, 2021. https://doi.org/10.48550/ARXIV.2105.15203.

XIE, W.; CHU, G.; QIAN, Q.; YU, Y.; ZHAI, S.; CHEN, D.; WANG, N.; BAO, H.; ZHANG, G. Omnidirectional Dense SLAM for Back-to-back Fisheye Cameras. 2024 IEEE International Conference on Robotics and Automation (ICRA). **Proceedings**... In: 2024 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA). Yokohama, Japan: IEEE, 2024. https://doi.org/10.1109/ICRA57147.2024.10610351.

YANG, Y.; PAN, M.; TANG, D.; WANG, T.; YUE, Y.; LIU, T.; FU, M. MCOV-SLAM: A Multicamera Omnidirectional Visual SLAM System. **IEEE/ASME Transactions on Mechatronics**, p. 1–12, 2024. https://doi.org/10.1109/TMECH.2023.3348986.

YOGAMANI, S.; HUGHES, C.; HORGAN, J.; SISTU, G.; VARLEY, P.; O'DEA, D.; URICAR, M.; MILZ, S.; SIMON, M.; AMENDE, K.; WITT, C.; RASHED, H.; CHENNUPATI, S.; NAYAK, S.; MANSOOR, S.; PERROTON, X.; PEREZ, P. WoodScape: A multi-task, multi-camera fisheye dataset for autonomous driving. **arXiv**, 1905.01489, 2019. Available at: <https://arxiv.org/abs/1905.01489>. Accessed: 26 Apr. 2024.

YU, C.; LIU, Z.; LIU, X.; XIE, F.; YANG, Y.; WEI, Q.; QIAO, F. DS-SLAM: A Semantic Visual SLAM towards Dynamic Environments. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). **Proceedings**... In: 2018 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS). Madrid, Spain: IEEE, 2018. p. 1168–1174. https://doi.org/10.1109/IROS.2018.8593691.

ZHANG, W.; WANG, S.; DONG, X.; GUO, R.; HAALA, N. BAMF-SLAM: Bundle Adjusted Multi-Fisheye Visual-Inertial SLAM Using Recurrent Field Transforms. 2023 IEEE International Conference on Robotics and Automation (ICRA). **Proceedings**... In: 2023 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA). London, United Kingdom: IEEE, 2023. https://doi.org/10.1109/ICRA48891.2023.10160905.

ZHANG, Y.; HUANG, F. Panoramic Visual SLAM Technology for Spherical Images. **Sensors**, v. 21, n. 3, p. 705, 2021. https://doi.org/10.3390/s21030705.

ZHANG, Z.; REBECQ, H.; FORSTER, C.; SCARAMUZZA, D. Benefit of large field-of-view cameras for visual odometry. 2016 IEEE International Conference on Robotics and Automation (ICRA). **Proceedings**... In: 2016 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION (ICRA). Stockholm, Sweden: IEEE, 2016. https://doi.org/10.1109/ICRA.2016.7487210.

ZHAO, X.; HU, Q.; ZHANG, X.; WANG, H. An ORB-SLAM3 Autonomous Positioning and Orientation Approach using 360-degree Panoramic Video. 2022 29th International Conference on Geoinformatics. **Proceedings**... In: 2022 29TH INTERNATIONAL CONFERENCE ON GEOINFORMATICS. Beijing, China: IEEE, 2022. https://doi.org/10.1109/Geoinformatics57846.2022.9963855.

ZHOU, H.; ZOU, D.; PEI, L.; YING, R.; LIU, P.; YU, W. StructSLAM: Visual SLAM With Building Structure Lines. **IEEE Transactions on Vehicular Technology**, v. 64, n. 4, p. 1364–1375, 2015. https://doi.org/10.1109/TVT.2015.2388780.

ZLOT, R.; BOSSE, M. Efficient Large-scale Three-dimensional Mobile Mapping for Underground Mines. **Journal of Field Robotics**, v. 31, n. 5, p. 758–779, 2014. https://doi.org/10.1002/rob.21504.