

UNIVERSIDADE FEDERAL DO PARANÁ

CARLOS EDUARDO SETENARES MAGRIN

CREATING A MULTI-SENSOR DATASET IN HYBRID INDOOR AND OUTDOOR
ENVIRONMENTS TO ASSIST WHEELED MOBILE ROBOTS WITH LOCALIZATION
AND MAPPING USING PROPOSED COGNITIVE BLENDED SENSORS

CURITIBA

2025

CARLOS EDUARDO SETENARES KI MAGRIN

CREATING A MULTI-SENSOR DATASET IN HYBRID INDOOR AND OUTDOOR
ENVIRONMENTS TO ASSIST WHEELED MOBILE ROBOTS WITH LOCALIZATION
AND MAPPING USING PROPOSED COGNITIVE BLENDED SENSORS

Tese apresentada como requisito parcial à obtenção do grau de Doutor em Ciência da Computação no Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, da Universidade Federal do Paraná.

Área de concentração: *Ciência da Computação*.

Orientador: Eduardo Todt.

CURITIBA

2025

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

Magrin, Carlos Eduardo Setenareski

Creating a multi-sensor dataset in hybrid indoor and outdoor environments to assist wheeled mobile robots with localization and mapping using proposed cognitive blended sensors / Carlos Eduardo Setenareski Magrin. – Curitiba, 2025.

1 recurso on-line : PDF.

Tese (Doutorado) - Universidade Federal do Paraná, Setor de Exatas, Programa de Pós-Graduação em Informática.

Orientador: Eduardo Todt

1. Sensoriamento remoto. 2. Robôs móveis. 3. Robôs móveis – Sistemas de controle. I. Universidade Federal do Paraná. II. Programa de Pós-Graduação em Informática. III. Todt, Eduardo. IV. Título.

Bibliotecário: Douglas Lenon da Silva CRB-9/1892

ATA DE SESSÃO PÚBLICA DE DEFESA DE DOUTORADO PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIA DA COMPUTAÇÃO

No dia tres de novembro de dois mil e vinte e cinco às 08:30 horas, na sala AUDITÓRIO, DEPARTAMENTO DE INFORMÁTICA, foram instaladas as atividades pertinentes ao rito de defesa de tese do doutorando **CARLOS EDUARDO SETENARESKI MAGRIN**, intitulada: **Creating a multi-sensor dataset in Hybrid Indoor and Outdoor Environments to Assist Wheeled Mobile Robots with Localization and Mapping using proposed Cognitive Blended Sensors**, sob orientação do Prof. Dr. EDUARDO TODT. A Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação INFORMÁTICA da Universidade Federal do Paraná, foi constituída pelos seguintes Membros: EDUARDO TODT (UNIVERSIDADE FEDERAL DO PARANÁ), JOAO ALBERTO FABRO (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ), EDUARDO JAQUES SPINOSA (UNIVERSIDADE FEDERAL DO PARANÁ), ANDRE SCHNEIDER DE OLIVEIRA (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ). A presidência iniciou os ritos definidos pelo Colegiado do Programa e, após exarados os pareceres dos membros do comitê examinador e da respectiva contra argumentação, ocorreu a leitura do parecer final da banca examinadora, que decidiu pela APROVAÇÃO. Este resultado deverá ser homologado pelo Colegiado do programa, mediante o atendimento de todas as indicações e correções solicitadas pela banca dentro dos prazos regimentais definidos pelo programa. A outorga de título de doutor está condicionada ao atendimento de todos os requisitos e prazos determinados no regimento do Programa de Pós-Graduação. Nada mais havendo a tratar a presidência deu por encerrada a sessão, da qual eu, EDUARDO TODT, lavrei a presente ata, que vai assinada por mim e pelos demais membros da Comissão Examinadora.

CURITIBA, 03 de Novembro de 2025.

Assinatura Eletrônica

04/11/2025 11:11:05.0

EDUARDO TODT

Presidente da Banca Examinadora

Assinatura Eletrônica

04/11/2025 10:39:15.0

JOAO ALBERTO FABRO

Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO
PARANÁ)

Assinatura Eletrônica

04/11/2025 10:21:01.0

EDUARDO JAQUES SPINOSA

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

05/11/2025 07:20:45.0

ANDRE SCHNEIDER DE OLIVEIRA

Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO
PARANÁ)

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação INFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **CARLOS EDUARDO SETENARESKI MAGRIN**, intitulada: **Creating a multi-sensor dataset in Hybrid Indoor and Outdoor Environments to Assist Wheeled Mobile Robots with Localization and Mapping using proposed Cognitive Blended Sensors**, sob orientação do Prof. Dr. EDUARDO TODT, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 03 de Novembro de 2025.

Assinatura Eletrônica

04/11/2025 11:11:05.0

EDUARDO TODT

Presidente da Banca Examinadora

Assinatura Eletrônica

04/11/2025 10:39:15.0

JOAO ALBERTO FABRO

Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ)

Assinatura Eletrônica

04/11/2025 10:21:01.0

EDUARDO JAQUES SPINOSA

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

05/11/2025 07:20:45.0

ANDRE SCHNEIDER DE OLIVEIRA

Avaliador Externo (UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ)

AGRADECIMENTOS

Agradeço, em primeiro lugar, aos meus pais, pelo apoio incondicional, pelos ensinamentos e pelos valores transmitidos, que foram fundamentais para minha formação pessoal e acadêmica. À minha esposa e às minhas filhas, expresso minha mais profunda gratidão pela compreensão, paciência e incentivo constantes, especialmente nos momentos de maior exigência desta longa jornada.

Conciliar as responsabilidades familiares e profissionais com as demandas de um doutorado foi um desafio que exigiu determinação e resiliência, intensificado pelos impactos da pandemia, que trouxe incertezas e adiamentos, mas também lições de perseverança e superação.

Estendo meus agradecimentos aos colegas do Laboratório de Visão, Robótica e Imagem (VRI) pelo convívio, pela colaboração técnica e pelas discussões enriquecedoras que contribuíram significativamente para o amadurecimento deste trabalho. Sou igualmente grato pelo apoio e pelos recursos disponibilizados pelo Professor Wander da Cruz e pelo Departamento de Geomática, cuja colaboração foi essencial para a construção da base de dados e para o desenvolvimento desta pesquisa.

Por fim, registro meu sincero reconhecimento ao meu orientador, pela confiança, dedicação e valiosas contribuições científicas, que foram essenciais para a consolidação deste trabalho e para o meu crescimento como pesquisador.

“Don’t give up. Give hope.”
— **The Hardest Run**

RESUMO

Aprender sobre um ambiente enquanto nos movimentamos é uma das funções do sistema cognitivo humano. Percebemos o que está ao nosso redor por meio dos sentidos básicos e, a partir disso, tomamos decisões sobre as ações a serem executadas. Para identificar um lugar, o sistema cognitivo, por meio da atenção, seleciona estímulos sensoriais relevantes para compreender o ambiente. No entanto, fatores como o período do dia, a estação do ano ou mesmo a presença de determinados objetos podem alterar o peso atribuído a cada sentido. Robôs móveis autônomos, de modo análogo, utilizam sensores para perceber o ambiente em que se locomovem. Contudo, ao depender apenas de um tipo de sensor, não conseguem identificar uma mesma cena em diferentes posições, sob variações de iluminação ou na presença de objetos dinâmicos. Da mesma forma, utilizar um grande volume de dados provenientes de múltiplos sensores não garante, por si só, uma observação adequada para localização e navegação. Neste trabalho, apresentamos uma revisão sobre o processo cognitivo humano e os sistemas cognitivos que, considerados em uma arquitetura hierárquica, envolvem o ciclo percepção-ação. A partir desse estudo, foi proposto o método de combinação cognitiva de sensores (CBS – *Cognitive Blended Sensor*) como apoio ao mapeamento e à localização de robôs móveis. O modelo de combinação de sensores sugere conjuntos (SBS – *Sensor Blend Sets*) com diferentes características dentro do processo de fusão do CBS, capazes de colaborar com tarefas relacionadas ao estado do robô, movimentação, medição de distância, orientação, mapeamento e posicionamento. Alguns direcionamentos foram adotados neste trabalho com o objetivo de guiar a proposta e contribuir para a área de estudo, entre eles: a pesquisa sobre sistemas cognitivos humanos e sistemas cognitivos dinâmicos; o desenvolvimento do conceito de robô móvel gêmeo digital; a análise de *datasets* que utilizam mapeamento com múltiplos sensores; o estudo de plataformas e métodos de referência para ambientes internos e externos com validação de trajetória; experimentos com algoritmos de localização e mapeamento simultâneos; além da classificação e reconhecimento de objetos. Os resultados obtidos evidenciam o avanço da proposta por meio do método CBS-SBS, do desenvolvimento do robô móvel VRI4WD, dedicado à aplicação de múltiplos sensores em ambiente ROS, e, principalmente, da construção do *dataset* VRI4WD UFPR-MAP. Este *dataset* se destaca por sua especificidade, disponibiliza dados de múltiplos sensores de baixo custo com transição entre ambientes internos e externos no campus da UFPR, posicionamento geodésico preciso que permite a validação das trajetórias. Essa contribuição preenche uma lacuna relevante na literatura, oferecendo uma base experimental para os estudos de fusão de sensores em ambientes híbridos, auxiliando o avanço da pesquisa em robôs móveis autônomos.

Palavras-chave: Fusão Cognitiva de Sensores; Combinação Cognitiva de Sensores; Conjuntos de Combinação de Sensores; Mapeamento de Múltiplos Sensores; Robô Móvel.

ABSTRACT

Learning about an environment while moving through it is one of the fundamental functions of the human cognitive system. Based on this perception, we perceive our surroundings through basic senses and decide on the actions to be executed. To identify a place, the cognitive system, through attention, selects relevant sensory stimuli to interpret the environment. However, factors such as time of day, season, or the presence of specific objects can influence the weight given to each sense. Autonomous mobile robots, in a similar way, rely on sensors to perceive their environment. Yet, when depending solely on a single type of sensor, they are unable to identify the same scene under different positions, illumination changes, or dynamic objects. Likewise, processing large volumes of data from multiple sensors does not necessarily ensure adequate perception for localization and navigation. This work reviews the human cognitive process and cognitive systems, considered within a hierarchical architecture involving the perception–action cycle. Based on this study, the Cognitive Blended Sensor (CBS) method is proposed to support the mapping and localization of mobile robots. The sensor combination model introduces Sensor Blend Sets (SBS) with distinct characteristics within the CBS fusion process, contributing to tasks such as robot state estimation, movement, distance measurement, orientation, mapping, and positioning. Several directions were pursued in this work to guide the proposal and contribute to the field, including research on human and dynamic cognitive systems, development of the digital twin mobile robot concept, analysis of datasets employing multi-sensor mapping, evaluation of reference platforms and methods for indoor and outdoor trajectory validation, experiments with simultaneous localization and mapping algorithms, and object classification and recognition. The results demonstrate the advancement of the proposed approach through the CBS-SBS method, the development of the VRI4WD mobile robot for multi-sensor applications within the ROS environment, and, most notably, the construction of the VRI4WD UFPR-MAP multimodal dataset. This dataset stands out for its specificity, providing low-cost multi-sensor data with transitions between indoor and outdoor environments on the UFPR campus and precise geodetic positioning that enables trajectory validation. This contribution fills a significant gap in the literature by offering an experimental foundation for sensor fusion studies in hybrid environments, thereby supporting the advancement of research in autonomous mobile robotics.

Keywords: Cognitive Sensor Fusion; Cognitive Blended Sensors; Sensor Blend Sets; Multi-Sensor Mapping Dataset; Wheeled Mobile Robot.

LISTA DE FIGURAS

1.1	The PhD Research Roadmap outlines the evolutionary path of doctoral research through its key development stages and contributions. Publications are highlighted with a solid border, while other significant milestones and developments are indicated with a dashed border. The core contributions of this thesis are further denoted as Key Milestones, visually emphasized with a yellow star (★) on the map. This highlighting strategy directs the reader's attention to the high-impact points that define the success of this research.	25
2.1	Principles of an accelerometer and the sensor integrated into a module used in mobile robots. (a) Working principle of the mechanical accelerometer; (b) principle of MEMS accelerometer through oscillation; (c) An example of commercial MEMS accelerometer used in mobile robots (Siegwart et al., 2011)..	26
2.2	Phase relationship between channels of phase-quadrature incremental encoders. The observed phase relationship between channel A and channel B pulses is used to determine the direction of the rotation. The third channel for synchronization pulse generates a reference (index) pulse per revolution (Siegwart et al., 2011). .	27
2.3	MEMS gyroscope measures angular movement (pitch, roll, and yaw) using the Coriolis effect. The left shows an integrated circuit using gyroscope MEMS technology. The right shows MEMS gyroscope uses the Coriolis effect; the principles of the Coriolis effect consider a mass moving in the direction v . When an angular movement is applied (red arrow), the mass experiences a force in the direction of the yellow arrow due to the Coriolis effect. The resulting physical displacement is read on a MEMS gyroscope using a capacitive sensing interface (STMicroelectronics, 2024).	28
2.4	Optical proximity sensors. The opposed mode configuration relies on the target passage between the emitter and detector to interrupt the beam. The retroreflective mode of retroreflectors increases the effective range and simplifies alignment. Diffuse mode relies on energy reflected directly from the target surface. The convergent mode can be used to ascertain the approximate distance to an object. Adapted from (Everett, 1995).	29
2.5	The Multipath Effect. The figure shows the multipath effect where the signal reaches the receiver antenna via multiple paths. This occurs due to the reception of both the direct signal from the satellite and various reflected signals (specular and diffuse) from surrounding surfaces, causing interference at the receiver. Adapted from (Souza, 2004)..	30
2.6	Laser Scanning Rangefinder. The left shows an outdoor LiDAR detection range of 30 m for intelligent robots (HOKUYO UST-30LX). The right shows LiDAR sensor mapping in a wide detection angle of 270°. Adapted from (HOKUYO, 2024).	31

2.7	The magnetoresistive effect in <i>permalloy</i> . This is assumed to be the x-direction and the current flow direction. A magnetoresistive sensor now relies on two basic effects: The strip resistance depends on the angle α between the direction of the current and the direction of the magnetization; the direction of magnetization and, therefore, α can be influenced by an external magnetic field H_y , where H_y is parallel to the strip plane and perpendicular to the preferred direction. When no external magnetic field is present, the <i>permalloy</i> has an internal magnetization vector parallel to the preferred direction (Stork, 2000).	33
2.8	Ultrasound HC-SR04. The modules include ultrasonic transmitters, receivers, and control circuits. The HC-SR04 module provides a 2 cm to 400 cm measurement function, and the ranging accuracy can reach 3 mm (ElecFreaks, 2004).	33
2.9	Typical intensity distribution of an ultrasound sensor. The figure shows the opening angle for the sound beam to obtain precise directional information about objects encountered. Ultrasound opening angle is a significant limitation since sound propagates in a conelike manner (Siegwart et al., 2011).	34
2.10	Three problems with ultrasound range readings: <i>foreshortening</i> , <i>specular reflection</i> , and <i>cross-talk</i> . <i>Specular reflection</i> - when the ultrasound transmitter signal cannot return directly to the receiver. <i>Crosstalk</i> - one ultrasound receives the signal sent by another close to it. <i>Foreshortening</i> - occurs in measurements more significant than the measuring range. Adapted from (Murphy, 2000).	35
2.11	Mobile robot model in virtual platform V-REP/CoppeliaSim and physical OCTO robot platform. Simulating a two-wheeled mobile robot with proximity sensors (ultrasound), yaw orientation taken from the robot pose in the scene (digital compass), and included objects for representing the access points (APs) in the scene, for simulated received signal strength (RSS) from a wireless network. . . .	37
2.12	The hierarchical paradigm method. In the cycle, the autonomous mobile robot perceives the environment (sense), perform (plan), and controll the motion (act). Adapted from (Murphy, 2000).	38
2.13	Mobile robots <i>see-think-act</i> cycle. The reference control scheme for mobile robot systems identifies much knowledge associated with mobile robotics, such as perception, motion control, localization, and path planning (Siegwart et al., 2011).	38
2.14	A general framework for the ACT production system, identifying the components and their interlinking processes with working memory. Declarative memory interlinks storage and retrieval, and production memory interlinks match and execution. Production memory feedback is the history of the application of existing productions. Adapted from (Anderson, 1995).	39
2.15	The knowledge-level system. The agent, perception of the environment, and action consider the goals and knowledge. The system represents the concept of hierarchical paradigm method used in the autonomous mobile robot, sense-plan-act cycle (Figure 2.12) and see-think-act (Figure 2.13) (Newell, 1994).	

2.16	The hierarchy describes a computer system as a system of electronic devices with electron physics laws, electrical circuits with Ohm's and Kirchhoff's laws, logic circuits with Boolean algebra, register-transfer systems with parallel logic, or programming systems with a sequential interpretation of programs in many knowledge levels. Adapted from (Newell, 1994)	42
2.17	Preparation versus deliberation trade-off. The space trade-off distributes limited development resources in a one-dimensional manner. The curves represent the AI systems, expert systems, and human tasks. Humans have better immediate knowledge (preparation), and computers (HITECH) have better search knowledge (Newell, 1994).	43
2.18	The total cognitive system: perception, cognition (long-term memory), and the motor system. The perceptual systems sense the environment and buffer it in working memory, encoding, parsing, and putting the information into the cognitive process. The decoding provides commands produced by the cognitive system, and working memory provides information available to internal components for use by the motor system. Adapted from (Newell, 1994)	43
2.19	Overview of the information-processing components in the executive-process interactive control (EPIC) architecture. The figure shows the human cognition model, interaction with the task environment, input in perceptual processors (auditory, visual, and tactile), output in motor processors (vocal, manual, and ocular), computer in cognitive processor production rules, and working memory consists of goals (Meyer and Kieras, 1997).	45
2.20	Cyclic operation of the cognitive system. The cyclic process involves the environment and connects the systems. Perception, the cognitive system perceives the environment; Memory, with the data changing over time, the contents of memory are subject to time constraints; Attention, as the computational resource is limited (memory), the attention system prioritizes resource allocation; Intelligence, the adaptive process continually adjusts the plan.	47
2.21	Perception–action cycle of a cognitive system. The sensory hierarchy observes the surrounding environment in three hierarchical levels: primary, unimodal, and polymodal. The motor hierarchy is an area dedicated to the system's actions, involving the cortex's primary, premotor, and prefrontal. Working memory represents a recent event for a pending action, and executive and perceptual memory include all knowledge acquired through a learning process. Adapted from (Haykin, 2012; Fuster, 2005).	47
2.22	Understanding ROS Nodes. The ROS MASTER is a service for ROS (<i>roscore</i>) is the first thing you should run when using ROS. A node is an executable that uses ROS to communicate with other nodes. Nodes can <i>publish</i> messages to a topic and <i>subscribe</i> to a topic to receive messages. Messages are ROS data used when <i>subscribing</i> or <i>publishing</i> to a topic. Adapted from OpenRobotics (2021).	50
2.23	Shows a real and virtual robot moving simultaneously. The architecture consists of the basic three-dimensional concept, a physical part (the UR3 robot), the virtual part (the model of the UR3 robot), and communication (Modbus communication protocol in Java application) (Pires et al., 2019).	51

2.24	Digital Twin Mobile Robot concept model. The <i>physical</i> system, a real mobile robot in a real environment; <i>virtual system</i> , mobile robot model in a virtual scene; <i>service</i> , integration between the ROS nodes; <i>data</i> , exchange of ROS messages by topics; <i>connection</i> , data exchange in cyber-physical systems through the ROS master (Magrin et al., 2021).	52
3.1	Sensors suite on the autonomous vehicle (VW Passat). The figure shows the multi-sensors in the vehicle, such as the inertial navigation system (GPS/IMU), laser scanner (Velodyne), and cameras (Point Grey Flea 2) (Geiger et al., 2013). .	58
3.2	Sensors suite on the Oxford RobotCar. The figure shows the multi-sensors in the vehicle, such as trinocular stereo camera (Bumblebee), monocular camera (Grasshopper 2), 2D LiDAR (LMS-151), 3D LiDAR (LD-MRS), and inertial/GPS navigation system (SPAN-CPT) (Maddern et al., 2017).	59
3.3	Sample imagery from the NCLT dataset. The figure shows samples in indoor and outdoor environments and at different times of the day (Carlevaris-Bianco et al., 2016).	60
3.4	Sensors suite on the M2DGR robot. The figure shows the multi-sensors in the wheeled mobile robot, such as 2 - monocular camera (FLIR Chameleon3), 3 - 3D laser scanner (Velodyne VLP-32C), 4 - GNSS receiver (Ublox M8T), 5 - IMU (Handsfree A9), 7 - infrared camera (Gaode PLUG 617), 8 - depth camera (Intel Realsense d435i), and 9 - event camera (Inivation DVXplorer) (Yin et al., 2022)..	62
3.5	Sensors suite on the TIERS robot. The figure shows the multi-sensors in the wheeled mobile robot, such as a 3D laser scanner (Velodyne VLP-16), 3D laser scanner (Ouster OS1-64), 3D laser scanner (Ouster OS0-128), 3D laser scanner (Livox Horizon), 3D laser scanner (Livox Avia), and LiDAR depth camera (Intel RealSense L515) (Qingqing et al., 2022).	62
4.1	Perception–action cycle of a cognitive sensor fusion. The sensory hierarchy observes the surrounding environment in three hierarchical levels: proprioceptive sensors, exteroceptive sensors, and sensor fusion. The motor hierarchy controls the mobile robot in the environment, involving locomotion, motor control, and path planning. Working memory represents the state of the mobile robot; actuator memory means the action from the motor hierarchy, and perception memory means the action from the sensory hierarchy through a sensor fusion process. Inspired from (Haykin, 2012; Fuster, 2005).	65
4.2	Sensor Blend Sets (SBS). Each of the five levels has a feature that combines and complements the process of a cognitive system. SBS-State (battery current, gyroscope, and bumper); SBS-Movement (accelerometer and encoder); SBS-Distance (ultrasound, infrared, and digital compass); SBS-Mapping (laser scanner and vision); and SBS-Positioning (GPS and RF sensor).. . . .	67
4.3	Learning perception-action cycle of cognitive blended sensors in each level. The figure shows the hierarchical levels of sensory and motor, such as environment, proprioceptive sensors and locomotion (level 0), exteroceptive sensors and motor control (level 1), sensor fusion and path planning (level 2 - SBS), and cognitive learning (attention) selection of the best sensor blend sets.	69

5.1	Maze environment of the Autonomous Mobile Robots Championship organized by the Laboratory ROSIE/UNICURITIBA. We used it in the design of the scene in Gazebo.	71
5.2	The Gmapping SLAM package shows in RViz and simulates locomotion in a maze Gazebo scene. Using a differential drive mobile robot in Gazebo, a 2-D occupancy grid map from the pose data and laser scanner mounted on top of the mobile robot was created.	71
5.3	The image map, which has the encoded data of the occupancy grid map, is saved using the map server package. The map server package stored the YAML file containing the map metadata and the image PGM file.	72
5.4	The physical TUPY-4WD robot was developed to apply multi-sensors and locomotion in indoor environments and is set to include and adjust many types of sensors.	72
5.5	The virtual TUPY-4WD robot in CoppeliaSim. The figure shows mounted on the top of the mobile robot eight ultrasound sensors in an octagon, a Kinect, and a Hokuyo laser scanner (Magrin et al., 2021).	73
5.6	Pinheirinho campus at the University of Curitiba. We used the floor plan to model indoor and outdoor environments in CoppeliaSim. The figure shows the vehicle parking and entrances 'C' block and 'D' block.	73
5.7	CoppeliaSim indoor and outdoor virtual environment of the Pinheirinho Campus at the University of Curitiba. Simulation TUPY-4WD robot model for locomotion and perception of the indoor environment, using a laser scanner and depth image/RGB image from Kinect.	74
5.8	Displaying graphs of running ROS nodes (ellipses) with connecting topics (rectangles) and packages (obtained by the tool <i>rqt_graph</i>). The ROS nodes running robot teleoperation <i>/csf_robotCS_teleop</i> , interface ROS with CoppeliaSim <i>/sim_ros_interface</i> , edge detection algorithm and OpenCV function with Kinect <i>/csf_robotCS_cannyKinect</i> , and preprocessing sensor information <i>/csf_robotCS_perception</i>	75
5.9	CoppeliaSim simulator with ROS interface running ROS packages for the perception of the TUPY-4WD model in the Pinheirinho environment. The figure shows laser scanning the scene, Kinect showing depth image/RGB image and edge detector output, ultrasound octagon reading distance, access point to the robot distance, and robot orientation.	76
5.10	Sample images from the NCLT dataset, session 2013-04-05. The figure shows images from places, such as research labs, central aisles, streets, and sidewalks (Carlevaris-Bianco et al., 2016).	77
5.11	Running the model in the Detectron2 library, images show object recognition in the samples of the NCLT dataset. The figure shows object recognition, such as a person, a backpack, a chair, a table, and a car.	78

5.12	Sensors suite on the VRI4WD Mobile Robot Platform. The figure shows multi-sensors in the robot, such as laser range scanner RPLIDAR A1, ultrasonic distance sensor DFRobot URM37, IMU with absolute orientation sensor Bosch BNO055, camera Microsoft Kinect XBOX 360, infrared DFRobot Sharp GP2Y0A02YK, quadrature encoders embedded in the front motors, and GPS u-blox M8 GNSS fixed on the top.	79
5.13	VRI4WD Robot hardware block diagram. At the top - are three blocks, and at the bottom - are two blocks representing the DC-regulated power for supplying the motors, electronic boards, and sensors for each voltage level. Light gray blocks represent the 8-bit microcontrollers, input analog and digital sensor data, DC motor driver output controls, and UART connection with the Raspberry Pi boards using a logic-level converter. Sensors were identified by colors representing SBS with different features.	80
5.14	VRI4WD Mobile Robot Platform. The mechanical platform is a 4WD all-terrain chassis. (a) The Figure shows the robot moving along the trajectory in an indoor environment. (b) The Figure shows the robot moving along the trajectory in an outdoor environment.	81
5.15	VRI4WD robot with Wireless D-Link DIR-600 2.4GHz IEEE 802.11b/g router embedded on the rear of the platform for more accessible communication with notebooks via ROS. The wireless router has a 5V DC / 1A power connector and an external fixed 5dBi antenna.	82
5.16	The position of the sensors at different levels ensures better recognition of the environment, ease of adjustment, and better validation of the other sensors. We balanced the robot's center of gravity with the position of the batteries at the bottom and the router at the rear. A bumper was mounted on the front of the robot to prevent it from tipping over while moving on uneven terrain.	84
5.17	Sensors are positioned at different levels on the VRI4WD robot platform. At the top, the laser scanner (RPLIDAR A1) and five ultrasound sensors.	84
5.18	The infrared sensors are at the same level as the Kinect and positioned front, left, and right to measure the distance in centimeters to an object.	85
5.19	The IMU BNO055 was on the last platform close to the center of the robot. The module was on a breadboard, as it was necessary to test different models. The BNO055 uses UART communication, converter module USB-UART (CP2102).	85
5.20	An adapter is needed to attach the VRI mobile robot's 4WD suspension. The figures show the adapter with height adjustment using upper screws and a clamp for attachment to the motor shaft, locking the suspension and preventing unnecessary platform swinging.	87
6.1	Relative Method. The coordinates of the unknown point "B" are calculated using the known coordinates of point "A" as a base. Adapted from (SUCI, 2012). . . .	89
6.2	GNSS and baselines. GNSS receiver over point RE01 was installed on a 2.00 m antenna, and an image from the processing software showed the vectors that were part of the baselines.	90

6.3	Example of a framed polygonal, starting A1 and A2 points and ending A3 and A4 points with known coordinates. P1 and P2 represent the pole (prism) to reference the positioning (Veiga et al., 2007)..	90
6.4	Support polygonal and points for validation. The support polygon, along with the names of the points, is presented in yellow, and the points used to validate this work are represented in red.. . . .	90
6.5	Measurement of each point in the building of the Geology Department in Engineering Street from Polytechnic Center - UFPR. (a) The procedure with the total station, using the pole (prism) to reference the landmark points. (b) Total station Leica TCR 407 from LAIG/LABTOP UFPR for the measurement of each point. (c) The figure shows the procedure for tying the outdoor points with the indoor ones made with the total station. (d) The method for tying the indoor points at the building.	91
6.6	VRI4WD UFPR-MAP, positions in a Google Earth map. The numbers indicate the 123 landmark points on the Polytechnic Center - Geology Building. It alphabetically represents the point-to-point trajectory of the robot on the map, indoor (AB and BF), and outdoor (BC, CD, DE, and EB).. . . .	92
6.7	Tags manufacturing in a printer, liquid resin. Showing the name of the dataset and the position number on the map. Tags with a diameter of 30 mm and a thickness of 2 mm, using a projection for the letters and a recess for the number. Used in two colors, gray or white, to differentiate multiples of 5 from white tags.	92
6.8	Tags fixed to the wall and pillars around the Geology building at the Federal University of Paraná - Polytechnic Center. The tags were positioned at an approximate height of 200 mm from the floor and tried to follow the architecture of the building to place the tags on each column or pillar, representing an approximate distance of 3 meters between the tags.	93
6.9	The environment used to build the database during the mobile robot's locomotion. Different surfaces and architectures were considered along the path, outdoor environment: (a) similar architecture, long distance for sensor measurement, and linear trajectory; (b) different architectures, parking, and irregular trajectory; (c) proximity between the walls on both sides and transition with the indoor environment; (d) different architectures, parking, and irregular trajectory; indoor environment, (e) narrow aisle for linear trajectory and measurement with left and right side sensors. Figure (f) shows the easy transition between indoor and outdoor environments through two side doors.	95
6.10	VRI4WD robot exploring indoor and outdoor UFPR-MAP. (a) The notebook was used to read the sensor topics through Wi-Fi communication embedded in the robot. (b) A robot is exploring the indoor environment. (c) A robot is exploring the outdoor environment. (d) Move the robot around the map with joystick. . . .	96
6.11	The RViz 3D visualization tool shows the setup of the relationships between coordinate frames in a tree structure. The TF VRI4WD platform indicated by the <i>base_footprint</i> and <i>base_link</i> , and the sensors relationship by the frames <i>ir_link</i> (3x infrared Sharp), <i>camera_link</i> (Kinect), <i>ultrasound_link</i> (5x URM37), <i>imu</i> (BNO055), and <i>laser</i> (laser scanner).. . . .	97

6.12	VRI4WD UFPR-MAP, paths taken on the map, Google Earth. Represents alphabetically the point-to-point trajectory of the robot in the map, indoor environment (AB and BF - <i>yellow path</i>), and outdoor environment (BC, CD, DE, and EB - <i>red path</i>).	98
7.1	VRI4WD proprioceptive sensors. The figure shows sensors reading data from encoders Left and Right in tick count and proprioceptive data from IMU: accelerometer (x, y, and z) and gyroscope (x, y, and z).. . . .	101
7.2	Odometry message. The figure shows the topic <code>/odom_data_quat</code> where orientation is quaternion.	101
7.3	The figure shows in RViz the VRI4WD trajectory axes obtained from the raw encoder data using the odometry package, considering the <code>/odom_data_quat</code> topic.	102
7.4	VRI4WD exteroceptive sensors. The figure shows sensors reading data from IR and ultrasound (Left, Front, and Right) in centimeters, latitude and longitude from GNSS, orientation in degree, and RF sensors from access points available in the UFPR campus. The dataset presents the latitude and longitude from and ground-truth (landmark) and point in the map for reference.. . . .	103
7.5	Camera using Kinect. The image on the left shows the topic <code>/camera/rgb/image_mono</code> , and on the right, the output with image processing using the OpenCV Canny Edge Detection algorithm.	103
7.6	Laser scan mapping. The image shows in RViz screen laser scan topic <code>/scan</code> from <i>AB</i> path of the dataset. VRI4WD Robot using laser scanner Slamtec RPLIDAR-A1 with 12m range in 360°.	104
7.7	IMU frame orientation. The image shows the fixed frame orientation in RViz screen Imu topic <code>/imu/data</code> . The axes of reference to Euler angles are roll (red axis), pitch (blue axis), and yaw (green axis).	105
7.8	2D plots of the outdoor path D to E and outdoor path E to B. The image shows the ground-truth (landmark) in red and the robot positioning (GNSS) in green, considering the longitude on the x-axis and the latitude on the y-axis.. . . .	106
7.9	2D plots of the outdoor path B to C and outdoor path C to D. The image shows the ground-truth (landmark) in red and the robot positioning (GNSS) in green, considering the longitude on the x-axis and the latitude on the y-axis.. . . .	107
7.10	2D plots of the indoor path B to F and outdoor path A to B. The image shows the ground-truth (landmark) in red and the robot positioning (GNSS) in green, considering the longitude on the x-axis and the latitude on the y-axis.. . . .	108
7.11	Afternoon walk in the outdoor paths using Garmin Forerunner 245 watch. Google Earth map shows the landmark points in blue, and the GPX format file plots the outdoor trajectory in red.	109
7.12	Afternoon walk in the indoor paths using Garmin Forerunner 245 watch. Google Earth map shows the landmark points in blue, and the GPX format file plots the indoor trajectory in yellow.	109

7.13	Image classification indoor or outdoor environment. The image shows the following paths in each environment; they are (a) indoor path-AB, (b) indoor path-BF, (c) outdoor path-BC, (d) outdoor path-CD, (e) outdoor path-DE, (f) outdoor path-EB.	110
7.14	The image shows the 2-D occupancy grid map generated by the GMapping SLAM package, visualized in RViz. The map was built from odometry and laser scanner data collected on the PATH-AB of the UFPR-MAP dataset.	113
7.15	The image shows the 2-D occupancy grid map generated by the GMapping SLAM package, visualized in RViz. The map was built from odometry and laser scanner data collected on the PATH-BF of the UFPR-MAP dataset.	113
7.16	The image shows the 2-D occupancy grid map generated by the GMapping SLAM package, visualized in RViz. The map was built from odometry and laser scanner data collected on the PATH-EB of the UFPR-MAP dataset.	114
7.17	The image highlights a failure of GMapping to maintain localization consistency during the trajectory, resulting in an incomplete and distorted map reconstruction. The map, visualized in RViz, was built using the GMapping algorithm from odometry and laser scanner data collected on the PATH-CD of the UFPR-MAP dataset.	114

LISTA DE ACRÔNIMOS

4WD	Four Wheel Drive
ACT	Adaptive Control of Thought
ACT-R	Adaptive Control of Thought – Rational
ANN	Artificial Neural Network
AP	Access Point
CAD	Computer-Aided design
CBS	Cognitive Blended Sensors
CDS	Cognitive Dynamic System
CPR	Counts Per Revolution
CST	Cognitive Systems Toolkit
DC	Direct Current
DT	Digital Twin
DTMR	Digital Twin Mobile Robot
EKF	Extended Kalman Filter
EPIC	Executive-Process Interactive Control
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HSF	Hierarchical Sensor Fusion
IMM	Interacting Multiple Model
IMU	Inertial Measurement Unit
KF	Kalman Filter
KNN	K-Nearest Neighbors Algorithm
LED	Light Emitting Diode
LAIG	Geodetic Instrumentation Laboratory
LiDAR	Light Detection and Ranging
MEMS	Micro Electro Mechanical Systems
MLP	Multilayer Perceptron
NASA	National Aeronautics and Space Administration
NCLT	North Campus Long-Term
OSC	Operational Space Control
PID	Proportional-Integral-Derivative
PRP	Psychological Refractory Period
PWM	Pulse Width Modulation
RADAR	Radio Detection and Ranging
RBPF	Rao-Blackwellized Particle Filter

RF	Radio Frequency
ROS	Robot Operating System
RL	Reinforcement Learning
SBS	Sensor Blend Sets
STL	Standard Triangle Language
RBMS	Brazilian Network for Continuous Monitoring of Systems
RFID	Radio Frequency Identification
RGB	Red-Green-Blue
RSS	Received Signal Strength
RSSI	Received Signal Strength Indication
SLAM	Simultaneous Localization and Mapping
SMS	Systematic Mapping Study
UFPR	Federal University of Paraná
UKF	Unscented Kalman Filter
URDF	Unified Robot Description Format
UTM	Universal Transverse Mercator
WLAN	Wireless Local Area Network
WMR	Wheeled Mobile Robot
YOLO	You Only Look Once

SUMÁRIO

1	INTRODUCTION	20
1.1	CHALLENGES AND MOTIVATION	23
1.2	OBJECTIVES AND CONTRIBUTION	24
1.3	PHD JOURNEY: MILESTONES & PUBLICATIONS	24
2	THEORETICAL BACKGROUND.	26
2.1	MOBILE ROBOT SENSORS	26
2.1.1	Proprioceptive sensors	26
2.1.2	Exteroceptive sensors	28
2.2	SENSOR FUSION	35
2.3	HIERARCHICAL SENSOR FUSION	36
2.4	THE ARCHITECTURE OF THE MIND.	38
2.4.1	The total cognitive system	42
2.5	COGNITIVE SYSTEMS	46
2.5.1	Perception-action cycle	46
2.5.2	Memory	48
2.5.3	Attention	48
2.5.4	Intelligence	49
2.6	ROBOT OPERATING SYSTEM (ROS)	49
2.7	DIGITAL TWIN	50
2.7.1	Digital Twin Mobile Robot	51
2.8	CHAPTER CONSIDERATIONS.	53
3	RELATED WORK	55
3.1	RESEARCH DIRECTIONS	55
3.2	DATASETS	57
3.3	CHAPTER CONSIDERATIONS.	62
4	THE PERCEPTION-ACTION CYCLE	64
4.1	COGNITIVE SENSOR FUSION.	64
4.1.1	Perception-action cycle	64
4.1.2	Memory	65
4.1.3	Attention	66
4.2	SENSOR BLEND SETS	66
4.3	COGNITIVE BLENDED SENSORS METHOD	67
4.4	CHAPTER CONSIDERATIONS.	68

5	MOBILE ROBOT PLATFORM	70
5.1	EARLY EXPERIMENTS.	70
5.1.1	Simultaneous Localization and Mapping	70
5.1.2	Mobile Robot Model - TUPY 4WD	70
5.1.3	Creating a Scene - Pinheirinho Environment	72
5.1.4	Locomotion and Perception - ROS interface	74
5.1.5	Classification and Object Recognition - NCLT Dataset	76
5.2	VRI4WD MULTI-SENSOR MOBILE ROBOT	77
5.3	CHAPTER CONSIDERATIONS.	87
6	BUILDING A DATASET	88
6.1	BUILDING THE VRI4WD UFPR-MAP DATASET	88
6.1.1	Creating a ground-truth	88
6.1.2	UFPR WLAN map	93
6.1.3	Sensor mapping dataset	94
6.1.4	UFPR-MAP dataset	95
6.2	CHAPTER CONSIDERATIONS.	98
7	EXPERIMENTS.	100
7.1	THE DIRECTION OF EXPERIMENTS	100
7.1.1	CBS Perception-Action Cycle.	100
7.2	RESULTS OF EXPERIMENTS	104
7.2.1	Landmarks <i>versus</i> GNSS	104
7.2.2	Environment Classification	105
7.2.3	Simultaneous Localization and Mapping	112
7.3	CHAPTER CONSIDERATIONS.	115
8	CONCLUSION AND FUTURE WORK	116
	REFERÊNCIAS	118

1 INTRODUCTION

Perception of the environment through multi-sensors has been widely addressed in the literature, employing approaches based on artificial intelligence and probabilistic methods such as the Kalman Filter for sensor fusion. The problem of localization and mapping in autonomous mobile robots, commonly referred to as Simultaneous Localization and Mapping (SLAM), relies on the integration of different types of sensors, including LiDAR, vision systems, and inertial measurement units (IMUs), for applications in both indoor and outdoor environments.

Moreover, autonomous mobile robots must perform tasks at different cognitive and operational levels, such as verifying basic conditions, whether the robot is on level ground, monitoring for collisions, or assessing battery status, before initiating locomotion. Closed-loop motion control depends on sensors providing odometry and/or acceleration measurements. Distance estimation in challenging environments, such as those involving translucent obstacles or poor lighting, can be improved using ultrasonic sensors. In outdoor environments, positioning information is typically obtained through GNSS systems. However, in indoor scenarios, satellite signals are often unavailable or highly degraded, thus requiring alternative positioning techniques such as RF sensor-based localization methods.

Advances in cognitive systems research have highlighted the potential of hierarchical sensor fusion architectures and multi-sensor combination models to enhance robotic perception and decision-making processes. These approaches, inspired by cognitive systems, address the perception-action loop in autonomous mobile robots by assigning different sensors to distinct tasks according to their suitability and reliability in varied contexts (Chen et al., 2020; Lange, 2019).

This research emphasizes the need for a comprehensive multi-sensor dataset to support the proposed Cognitive Blended Sensors methodology. Consequently, a dedicated dataset was developed, integrating various types of sensors on a mobile robotic platform designed specifically for sensor mapping in indoor and outdoor environments. This dataset significantly contributes to the field by providing diverse sensor data to advance sensor fusion, localization, and mapping research.

For a dynamic system to be cognitive, it has to embody five distinct processes: perception, memory, attention, language, and intelligence (Fuster, 2005). Haykin (2012) proposed a new class of dynamic systems called cognitive dynamic systems, based on Fuster (2005) work, inspired by human cognition like a form of computation. In other words, the dynamic system operating in an environment to be explored is cognitive if it can have four fundamental functions basic to human cognition: the perception-action cycle, memory, attention, and intelligence. We leave out language because it is outside the scope of the cognitive dynamic system.

Lange (2019) predicts that the 2020s will be the decade when progress in cognitive abilities, that is, the ability to sense, understand, and interact, will significantly improve robotics applications. Cognitive robotics provides robots with cognitive skills similar to humans and animals. Acquiring knowledge through actions or perceptions is an essential focus of cognitive robotics research. For teaching people skills, robots are trained to include the ability to sense, understand, and interact. The autonomous robot is then trained to perform a task by learning a mapping between observations and actions. Simulation, reinforcement learning, and computer vision will likely provide robots with basic cognitive abilities.

The faster development of robotic hardware, including cameras and LiDAR, reduces the cost of acquisition, ownership, and maintenance, combined with ease of use, will accelerate

the growth of cognitive robotics (Lange, 2019). In addition, Chen et al. (2020) description of deep learning motivated researchers to consider data-driven methods as an alternative to solve problems in real and dynamic environments with imperfect sensor measurements and inaccurate system modeling. Conventional solutions are achieved by hand-designing algorithms and calibrating them to a particular application domain, whereas learning-based approaches construct this mapping function by learning from knowledge. Learning techniques rely on massive datasets, which are computationally more costly, to extract statistically meaningful patterns and can struggle to generalize to out-of-sample environments. Chen et al. (2020) categorize deep learning for localization and mapping approaches into *odometry estimation* - concerning the calculation of the relative change in pose, in terms of position and orientation, between two or more frames of sensor data; *mapping* - builds and reconstructs a consistent model to describe the surrounding environment; *global localization* - retrieves the global pose of a mobile robot in a known scene with prior knowledge; *SLAM* - optimizes odometry estimation, global localization and mapping to boost performance in both localization and mapping.

Localization and mapping are fundamental needs for human and mobile agents. *Localization* refers to the ability to obtain the internal system states of robot motion, including locations, orientations, and velocities. *Mapping* indicates the capacity to perceive external environment states and capture the surroundings, including geometry, appearance, and semantics of a 2D or 3D scene. Humans can perceive their self-motion and environment using their multi-modal perception, developing cognition to locate and navigate themselves in a complex three-dimensional space. Similarly, mobile robots should also perceive the environment and estimate their system state using onboard sensors, sensing their surroundings, and autonomously making decisions. An autonomous mobile robot requires precise and robust Localization to process new information and adapt to various scenarios continuously (Chen et al., 2020).

Considering the advancement of learning algorithms focused on efficiently learning from data received from the environment, we need to be concerned about how that data is perceived in the first place. The cognitive filter that protects humans from sensory overload is called attention. Reinforcement learning can be implemented by a dynamic and state-responsive mechanism, capable of acquiring behavioral adaptations to the environment characteristics during the learning process (Ramicic and Bonarini, 2020).

According to Chen et al. (2020), beyond the common choice of onboard sensors, such as cameras, IMU, and LiDAR, the emerging new sensors provide an alternative to constructing a more accurate and robust multimodal system. New sensors, including neuromorphic vision sensors (event cameras) (IniVation, 2020), thermo cameras, mmWave radar sensors (Iovescu and Rao, 2017), radio signals, and magnetic sensors, have distinct properties and data formats compared to predominant SLAM sensors such as cameras, IMU, and LIDAR. Nevertheless, the practical learning approaches to processing these unusual sensors are still underexplored.

A systematic mapping study on multi-sensor fusion (Magrin et al., 2019) identifies the main methods of sensor fusion and sensor types for application in mobile robot self-localization.

The study answered the following research questions over the last thirty years:

- *What are the main sensors used for WMR (Wheeled Mobile Robot) localization, and which configurations with multi-sensors are most applied?* The main sensors used are an accelerometer, digital compass, encoder, GPS, gyroscope, IMU, infrared, laser, RF sensor, ultrasound, and vision. The trend is to apply vision, ultrasound, laser, and encoder as the primary sensors used in multi-sensor fusion systems. The study found that there was a significant trend in using laser, GPS, and IMU applications. The main configurations of multi-sensors represent the tendency to apply vision and laser in the fusion of multi-sensors. In addition, the study shows that most works use only two types

of sensors, so the most significant representation with more than two types of sensors was with vision, laser, and encoder.

- *Which multi-sensor fusion methods are used for WMR localization, and which method represents the state-of-the-art?* The multi-sensor fusion methods are Bayes-estimation, behavior-based, Dempster-Shafer, fuzzy logic, genetic algorithm, Hough transform, Kalman filter, Markov, Monte Carlo, neural networks, particle filters, probabilistic, Rough set, rule-based, and weighted average. A significant trend using the Kalman Filter (KF) included the Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF) as sensor fusion methods for WMR localization. The research indicates advances in the representation of the KF as a sensor fusion method, pointing to the Kalman filter still being the state-of-the-art sensor fusion method applied.

In addition, other sensor fusion methods, such as neural networks, fuzzy logic, and particle filtering, are potential techniques for a hybrid fusion of multi-sensors. The study shows that the most widely used, empirically, sensor pairs are multi-sensor fusion using *vision and laser* and *ultrasound and encoder* for low-cost applications. The systematic mapping study shows a predominance in the use of two sensors (Magrin et al., 2019).

The survey of Lowry et al. (2016) provides an overview of the place recognition problem and its relationship with many robotics research fields, including localization, mapping, and recognition. The authors argue that visual place recognition is a highly challenging problem to solve in the general sense, given an image of a place. Defining a place recognition system must represent the environment (map) compared to the incoming visual data. Places in an environment may look very similar, and they may not always revisit places from the same viewpoint and position as before. The place recognition system must report a belief about whether or not the current visual information is from a place already included in the map. In mobile robotics, vision is the primary sensor for many localization and place recognition algorithms. Place recognition can also be made more efficient using hierarchical searching at the place level. Semantic context can support learning and predicting the changes in a scene and help increase robustness against environmental condition changes.

According to Lange (2019), robotics simulation is an emerging field; simulation in robotics cannot generate the quality and relevant training data needed to model the physical world properly. Generated scenarios may not lead to sufficient generalization for the robot to interact with humans effectively. The *Digital Twin* (DT) concept has been widely applied in industrial-relevant application fields, such as manufacturing, aerospace, electric power generation, automotive, oil and gas, healthcare and medicine, city management, and agriculture. Based on *digital twin* concept, Tao et al. (2019b) is considered an efficient way to realize the integration between physical and virtual spaces and represents the next surge of digitalization.

Recent reviews in autonomous systems research, such as the one by Xiang et al. (2023), emphasize the central role of multi-sensor fusion in enabling robust perception for autonomous driving. Their work presents a detailed taxonomy distinguishing symmetric and asymmetric fusion strategies across data, feature, and result levels, addressing the limitations of single-sensor modalities regarding robustness and spatial completeness. This classification framework clarifies existing algorithmic strategies and highlights the importance of selecting appropriate fusion schemes based on task requirements and environmental conditions. Although primarily developed in the context of autonomous vehicles, the fusion paradigms explored in their review are directly applicable to mobile robot navigation in hybrid environments, reinforcing the need for datasets and platforms that support the integration and benchmarking of diverse sensor modalities.

The recently addressed concept is multimodal sensor fusion, which combines information from various types of sensors to achieve a more complete and robust understanding of the environment. This technique is fundamental to improving accuracy and robustness by overcoming the individual limitations of different sensors. By exploiting various sensors' strengths, multimodal sensor fusion coherently estimates system states, such as a mobile agent's locations and orientations. This approach is required because many system state variables are not always fully observable by a single sensor modality. Consequently, it has become a central problem in robotics and computer vision, with wide-ranging applications in perception, planning, and control (Chen et al., 2025).

The construction of knowledge in the field of mobile robotics, particularly within hybrid environments, occurs through a balance between practical experimentation and theoretical methods. The proposed VRI4WD platform and the cognitive sensor fusion model developed in this thesis exemplify this dynamic. By integrating low-cost sensors and promoting the creation of datasets that capture the complexity of both indoor and outdoor environments, the work enables an active and adaptive learning process in which the robot not only perceives but also suggests methods to adapt to its environment, thereby enhancing its cognitive capabilities. This constructivist approach acknowledges that the development of efficient robotic systems depends on integrating the perception-action cycle, supported by an architecture that selectively assists in learning to combine sensory inputs according to situational demands. Consequently, this research contributes to the collaborative and incremental construction of knowledge in mobile robotics, encouraging future studies to explore hierarchical sensor fusion and the challenges encountered in complex environments, thereby advancing the development of autonomous mobile robots.

1.1 CHALLENGES AND MOTIVATION

Understanding and interacting with the environment remains a significant challenge for autonomous mobile robots. Sensor fusion involves discerning between disparate indoor and outdoor environments, maintaining robust perception without ideal lighting conditions, and accurately establishing the robot's precise position and orientation. The construction of reliable cognitive or mental maps, important for spatial learning, typically relies on the intelligent agent's consistent and highly corresponding observations of the environment. However, a fundamental difficulty arises because each type of sensor provides distinct information, and leaning on a single sensor modality often proves insufficient for comprehensive environmental understanding. Conversely, integrating diverse sensors can lead to overwhelming heterogeneous data, complicating the identification of the most relevant information for accurate environmental matching and state estimation.

This intrinsic complexity motivates a deeper investigation into how intelligent systems process various sensory inputs. Inspiration comes from studies on the human cognitive system, which reveal intricate interrelationships among key cognitive processes such as perception, memory, attention, language, and intelligence. Adapting these principles offers a compelling approach to address the complicated problems of mobile robot self-localization, particularly in challenging hybrid environments. The aim is to move beyond conventional sensor fusion methods by developing more adaptive and intelligent mechanisms to selectively process and combine sensory data, similar to cognitive functions, to achieve superior environmental awareness and ultimately enable more reliable autonomous navigation. This research, therefore, sets the stage for our objectives and contributions in advancing robust multi-sensor perception for mobile robots, under the sensor model and experimentation dataset perspectives.

1.2 OBJECTIVES AND CONTRIBUTION

The general objective of this work is to propose a cognitive blended sensor model as well as to create a suitable dataset to map the main sensors used in mobile robotics for hybrid indoor and outdoor environments. A sensor blend set concept is proposed to target the best sensors for each level of environmental features. The specific objectives of this work are to define and apply a cognitive method that considers different types of sensors and different hybrid environments to assist in the problem of self-localization and mapping for an autonomous mobile robot.

The specific objectives are as follows:

- the identification of cyclic operation of the cognitive system and the construction of an analogy between the perception-action cycle inspired by human cognition and sensor fusion for mobile robot self-localization;
- based on a cognitive process, propose a blended sensors method and also advances in indoor and outdoor sensor fusion using a sensor blend sets model;
- to build a mobile robot platform dedicated to this work and to create a multi-sensor mapping dataset in hybrid indoor and outdoor environments for blended sensor studies.

Based on the objectives, the proposal raises the following hypotheses:

- (H1) the concept of hierarchical sensor fusion can be extended to a cycle inspired by the human cognitive system;
- (H2) the cognitive blended sensors, effectively supported by sensor blend sets, significantly advance real-world localization and mapping by leveraging multi-sensor perception in hybrid indoor and outdoor environments for mobile robots;
- (H3) a hybrid indoor/outdoor dataset contributes to research in wheeled autonomous mobile robots, particularly if it includes features from multi-sensors necessary for blending sensors and ground-truth building with coordinate points obtained by geodetic and topographic methods.

1.3 PHD JOURNEY: MILESTONES & PUBLICATIONS

This section details the doctoral research roadmap, outlining the progression from the initial literature review to the final contributions, as visually represented in Figure 1.1. The research started with a comprehensive systematic mapping study on multi-sensor fusion for wheeled mobile robot self-localization, which critically evaluated the state-of-the-art and laid the foundation for subsequent methodological developments. Early advancements in this phase included a multi-sensor fusion method based on Artificial Neural Networks (ANN) and simulations of mobile robot localization based on Hierarchical Sensor Fusion (HSF). Initial mobile robot experiments complemented these theoretical and simulated efforts, effectively bridging the gap towards practical application.

The research progressed into developing key enabling tools and platforms, including building the VRI4WD mobile robot and extensive research about cognitive systems. A significant contribution to the field was the creation of a Digital Twin for Mobile Robots (DTMR) as an open-source learning tool. Subsequently, the work involved presenting the main multi-sensor datasets, supporting the development of a proposal for sensor blend sets and a cognitive blend

sensors method. These efforts resulted in the recent publication on multi-sensors for sensor blend sets in hybrid environments, showcasing the integration of the proposed methods. Ultimately, due to the lack of a multi-sensor dataset, the roadmap culminates in creating a geodetic ground truth and developing the VRI4WD UFPR-MAP for wheeled mobile robots in hybrid environments, ready to serve as a comprehensive resource for future research.

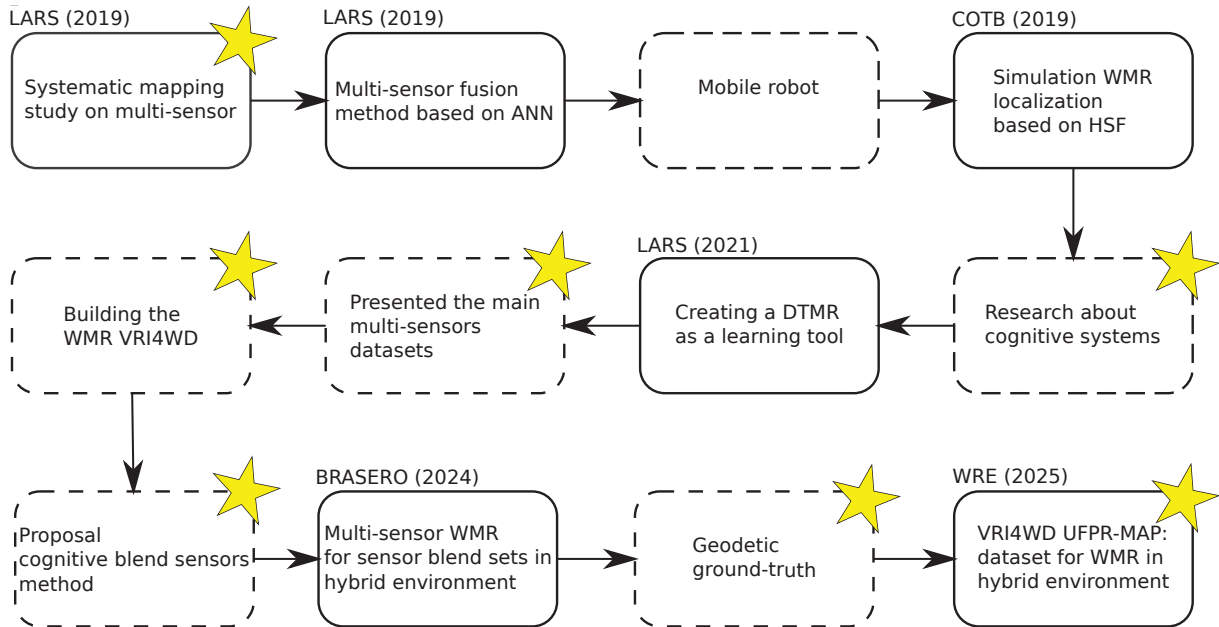


Figura 1.1: The PhD Research Roadmap outlines the evolutionary path of doctoral research through its key development stages and contributions. Publications are highlighted with a solid border, while other significant milestones and developments are indicated with a dashed border. The core contributions of this thesis are further denoted as Key Milestones, visually emphasized with a yellow star (★) on the map. This highlighting strategy directs the reader's attention to the high-impact points that define the success of this research.

2 THEORETICAL BACKGROUND

This chapter presents a literature review on the main sensors used for wheeled mobile robot (WMR) sensor fusion, multi-sensor fusion methods, and different data fusion architectures. We introduce hierarchical sensor fusion and the prominent publications on multi-sensor fusion published by the authors. Also, we reviewed the architecture of the mind to represent the perception-action system and the total cognitive system in psychology. Finally, we present the operation process of a cognitive system, the ROS framework, and the *digital twin* concept, which is covered as the base for our work.

2.1 MOBILE ROBOT SENSORS

The priori of this thesis involved a comprehensive review of the main sensors employed in wheeled mobile robots (WMR) for sensor fusion, conducted through a systematic mapping study published in LARS 2019 (Magrin and Todt, 2019a). For clarity of presentation, these sensors are classified into proprioceptive, which provide information on the robot's internal state, and exteroceptive, which capture data from the surrounding environment according to their established classification.

2.1.1 Proprioceptive sensors

Proprioceptive sensors monitor the robot's internal state, such as battery voltage, acceleration, and wheel movement. The main proprioceptive sensors used in WMR for sensor fusion are the accelerometer, encoder, and gyroscope.

2.1.1.1 Accelerometer

An accelerometer is a device used to measure acceleration along a single axis, with external forces acting upon it, including gravity projected on two to three orthogonal axes. Basically, an accelerometer is a *spring-mass-damper* system in which the three-dimensional position mass relative to the accelerometer can be measured in the working principle of a mechanical accelerometer. Figure 2.1 shows a basic accelerometer, Micro Electro Mechanical Systems (MEMS) accelerometer technology through oscillation, and a commercial accelerometer sensor in an integrated circuit (Sieewart et al., 2011).

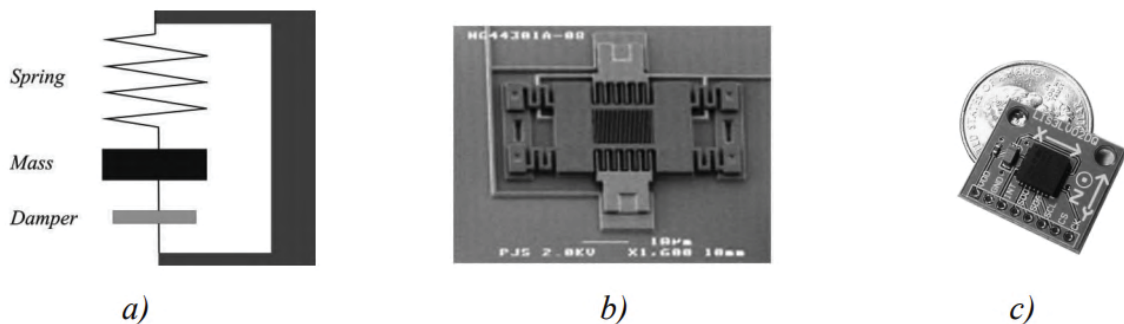


Figura 2.1: Principles of an accelerometer and the sensor integrated into a module used in mobile robots. (a) Working principle of the mechanical accelerometer; (b) principle of MEMS accelerometer through oscillation; (c) An example of commercial MEMS accelerometer used in mobile robots (Sieewart et al., 2011).

2.1.1.2 Encoder

The encoder is a device used to measure the position and speed of the robot's wheels and is responsible for measuring the internal and dynamic state of the mobile robot. The encoder is a sensor that converts angular or linear movement into digital signals, offering excellent resolution at a low cost (Siegwart et al., 2011). We can differentiate the encoder into incremental or absolute. Incremental encoders indicate the displacement about a previous position, measures the rotational speed, and allows estimating the relative position; it loses the position if the previous state is lost, while the absolute encoder indicates the displacement about a reference, generally using the *Gray code*, measures the angular position and allows you to estimate speed. The simplest encoder is the incremental type, which is a single-channel, perforated disc with the light emitter and receiver aligned with the holes in the disc; the control system can count the number of steps and determine the speed, but it does not resolve the direction of rotation of the robot. The phase-quadrature incremental encoder overcomes the problems by adding a second channel to better estimate the robot's movement in the environment. Figure 2.2 shows a phase-quadrature encoder. The encoder channels A and B pulses determine the direction of rotation. Here, the holes in the disc of channels A and B are 90 degrees out of phase. The lag between the channels allows you to identify the direction of rotation. Most encoders incorporate a third channel for synchronization pulse generated by the rotation of the disk, known as index (I) output (Everett, 1995).

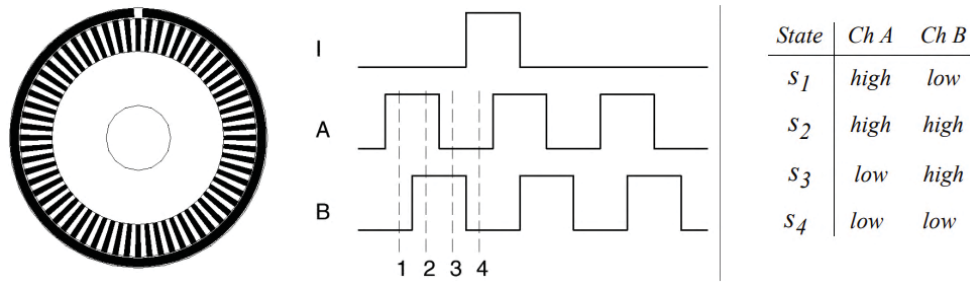


Figura 2.2: Phase relationship between channels of phase-quadrature incremental encoders. The observed phase relationship between channel A and channel B pulses is used to determine the direction of the rotation. The third channel for synchronization pulse generates a reference (index) pulse per revolution (Siegwart et al., 2011).

2.1.1.3 Gyroscope

Gyroscopes measure angular position by preserving their orientation with respect to a world coordinate frame, thus providing an absolute measurement for mobile systems. They are classified into two categories: mechanical gyroscopes and optical gyroscopes. The concept of a mechanical gyroscope depends on the inertial properties of a fast-spinning rotor. Optical gyroscopes are angular speed sensors that use two light beams emitted from the same source instead of moving mechanical parts (Siegwart et al., 2011). Commercial 3-axis gyroscopes are usually combined with an accelerometer in a standard package to allow advanced algorithms like sensor fusion. The gyroscopes used in mobile robotics are found in modules with an integrated circuit using MEMS technology. MEMS gyroscope measures angular movement (pitch, roll, and yaw) using the Coriolis effect - measures angular motion by detecting the displacement of vibrating structures caused by the interaction between the device's linear motion and rotational movement. It combines it with an accelerometer into an Inertial Measurement Unit (IMU) (STMicroelectronics, 2024).

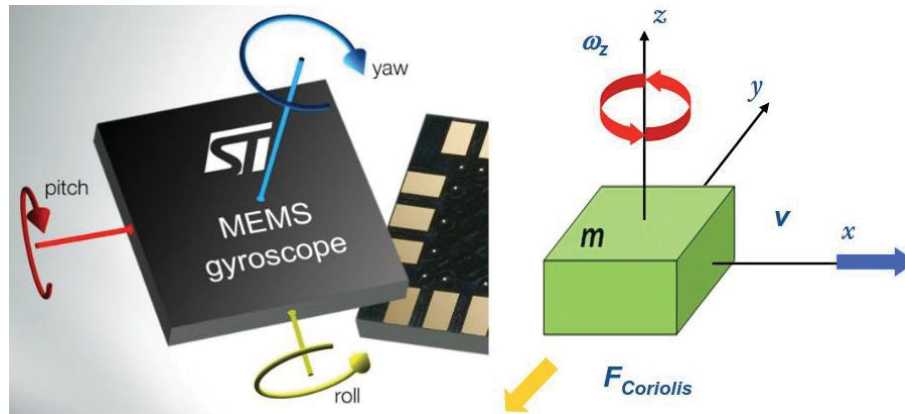


Figura 2.3: MEMS gyroscope measures angular movement (pitch, roll, and yaw) using the Coriolis effect. The left shows an integrated circuit using gyroscope MEMS technology. The right shows MEMS gyroscope uses the Coriolis effect; the principles of the Coriolis effect consider a mass moving in the direction v . When an angular movement is applied (red arrow), the mass experiences a force in the direction of the yellow arrow due to the Coriolis effect. The resulting physical displacement is read on a MEMS gyroscope using a capacitive sensing interface (STMicroelectronics, 2024).

2.1.2 Exteroceptive sensors

Exteroceptive sensors monitor the environment where the robot is located and capture information from the environment, such as proximity, distance, and images. The main exteroceptive sensors used in WMR for sensor fusion are infrared, GPS, laser, RF sensor, digital compass, ultrasound, and vision.

2.1.2.1 Infrared

Optical proximity sensors are based on the existence of an emitter (infrared) and a receiver (phototransistor or photodiode). We can divide proximity sensors into four groups according to application mode: opposed mode, retroreflective mode, diffuse mode, and convergent mode (Figure 2.4).

- *Opposed mode* the emitter and receiver are mounted separately. When aligned, they create a light barrier; any object that passes between the emitter and the receiver interrupts the light beam, opening the circuit (Everett, 1995).
- *Retroreflective mode* the emitter and receiver are mounted on the same encapsulation; a beam of light is sent from the emitter to the receiver through a reflector, and interruptions in the beam detect objects (Everett, 1995).
- *Diffuse mode* similar to retroreflective, the emitter and receiver are also assembled in the same encapsulation; an emitted beam of light creates a region where the presence of an object causes the diffusely reflected light to reach the receiver (Everett, 1995).
- *Convergent mode* configuration of the transmitter about the receiver ensures more accurate positioning information. The optical axis of the transmit LED (Light Emitting Diode) is tilted relative to the detector, so the two only intersect in a defined region - detection zone (Everett, 1995).

According to Siegwart et al. (2011), infrared sensors operate between 80 mm and 800 mm, a small distance compared to laser rangefinders and ultrasound sensors. The infrared sensor

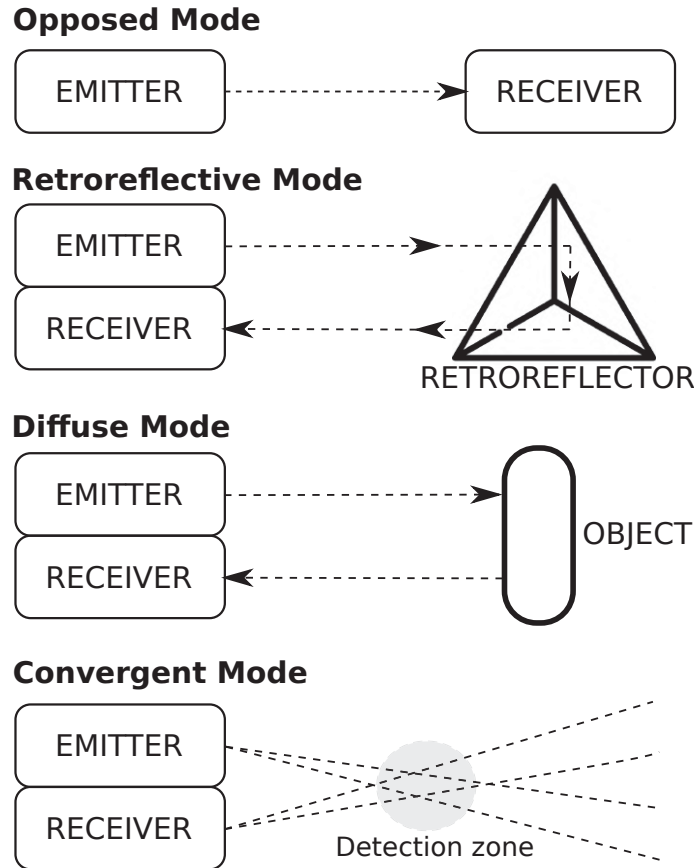


Figura 2.4: Optical proximity sensors. The opposed mode configuration relies on the target passage between the emitter and detector to interrupt the beam. The retroreflective mode of retroreflectors increases the effective range and simplifies alignment. Diffuse mode relies on energy reflected directly from the target surface. The convergent mode can be used to ascertain the approximate distance to an object. Adapted from (Everett, 1995).

has greater bandwidth and does not suffer from the cross-sensitivities common in ultrasound sensors, although it is more limited in range than ultrasound. Transparent objects may not be detected by infrared.

2.1.2.2 GNSS

The Global Navigation Satellite System (GNSS) allows users to obtain position, navigation, velocity, and time information through a receiver compatible with the system and capable of processing signals from positioning satellites. Positioning can be improved by combining signals from multiple systems, and there are currently four operational GNSS constellations: GPS (USA), GLONASS (Russia), BeiDou (China), and Galileo (European Union).

The Global Positioning System (GPS) was initially developed for military use but is now freely available for civilian navigation. Each satellite continuously transmits data that indicates its location and the current time. The receiver can infer its position by combining information regarding the arrival time and instantaneous location of four satellites. In confined spaces such as city blocks with tall buildings, or indoor spaces, the absence of sufficient sky visibility can prevent the GPS receiver to function adequately (Siegwart et al., 2011).

According to Siegwart et al. (2011), several factors affect the performance of a localization sensor that uses the GPS. First, it is essential to understand that because of the specific orbital paths of the GPS satellite coverage not being geometrically identical in different portions of the Earth, resolution is not uniform. Second, GPS satellites are merely an information source. They

can be employed with various strategies to achieve dramatically different levels of localization resolution.

Multipath occurs when the GNSS signal reflects off objects near the positioning and reaches the receiver via multiple paths. Due to the environmental features used to map the sensors, a complementary study was conducted on the multipath effect.

Multipath effect

According to Souza (2004), some effects influence the accuracy of GPS positioning. The main impact analyzed is multipath, where the quality of the signal is affected during its propagation between the satellite and the receiver antenna. In this case, the signal that reaches the GPS receiver antenna is the result not only of the signal coming directly from the satellite but also of secondary signals, resulting from the reflection of the signal on objects close to the antenna.

Figure 2.5 shows the multipath effect when the Global Navigation Satellite System (GNSS) signal reflects off objects near the positioning and reaches the receiver via multiple paths. Generally, the GNSS receiver receives, in addition to the direct signal, the reflected signal, which is delayed in relation to the direct signal. The multipath is caused mainly by signal reflections on surfaces close to the receiver, such as buildings, cars, and trees (de Souza, 2008).

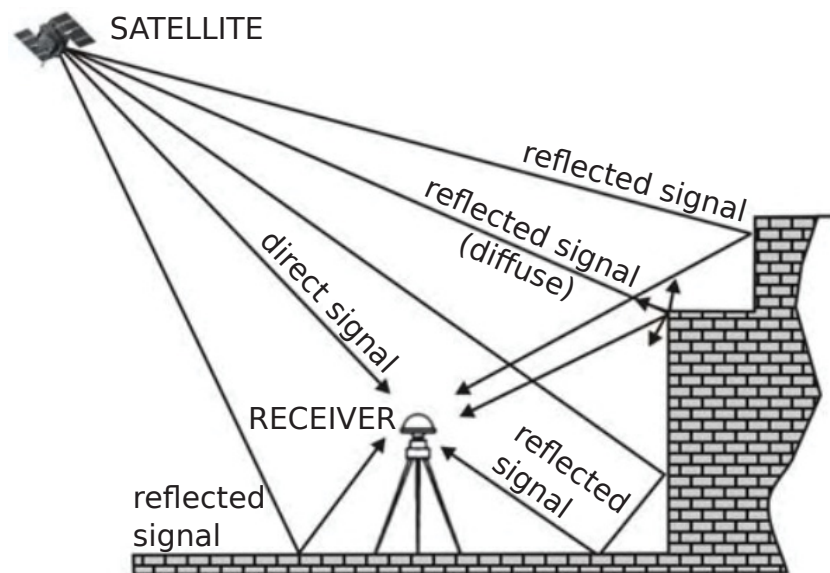


Figura 2.5: The Multipath Effect. The figure shows the multipath effect where the signal reaches the receiver antenna via multiple paths. This occurs due to the reception of both the direct signal from the satellite and various reflected signals (specular and diffuse) from surrounding surfaces, causing interference at the receiver. Adapted from (Souza, 2004).

Multipaths present errors in pseudo-distance measurements. The pseudo-distance measurement is obtained from the correlation between the code generated by the satellite at the time of transmission and its replica generated at the receiver at reception, depending on the scene's geometry involving antennas, satellites, and reflecting objects. It is essential to have means that can mitigate multipath effects to reduce the intensity of secondary signals and isolate the direct signal. Using special antennas, an arrangement of several antennas, an antenna location strategy, and observation of the signal over a long period is a technique that reduces the multipath effect (de Souza, 2008).

The GNSS technology is best suited for high-accuracy applications and outdoor environments. However, it is challenging to achieve high-accuracy positioning via smartphones or low-cost GNSS modules without a special antenna, in addition to the vulnerability of possible

antenna failures or even the position of the antennas, which could compromise the quality of positioning. Likewise, the estimated position can be obtained with WLAN technology, which is generally used to get the estimated position in indoor environments, as its estimated communication range is 100 meters, depending on the Wi-Fi standard, so it is necessary to have a structure with several access points for adequate correspondence (Gomes, 2023). The research of Gomes (2023) shows that absolute positioning accuracy with smartphones is 2 to 3 meters under good visibility conditions and satellite availability.

2.1.2.3 Laser rangefinder

Laser proximity sensors constitute optical sensors that measure range from centimeters to meters. They are commonly called LiDAR (Light Detection and Ranging). The distance is computed from the *time-of-flight*. Rotary lasers are based on the wave 360° rotation at more than 1 or 2 rpm and a mirror at 45°. They measure angular position and have problems with hidden areas. Scanning laser range combines the LiDAR with the rotary lasers by providing the detected object's range and angular position (Tzafestas, 2014).

The laser rangefinder is a *time-of-flight* sensor that follows the same principle as ultrasound. Furthermore, using light instead of sound significantly improves the range of the ultrasound sensor. However, unlike the ultrasound sensor, the laser rangefinder cannot detect the presence of transparent objects such as glass. This type of sensor consists of a transmitter that illuminates a target with a laser beam and a receiver that detects the light component that is essentially coaxial with the transmitted beam (Siegwart et al., 2011). A typical LiDAR for intelligent robots is shown in Figure 2.6, which detects the size, position, and moving direction of objects.

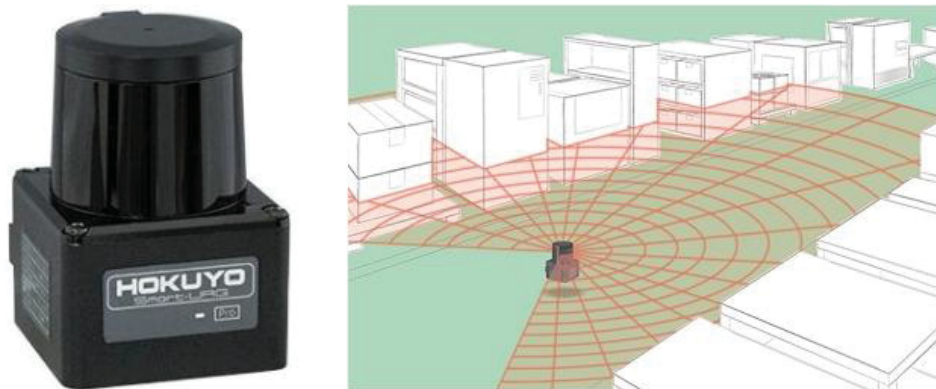


Figura 2.6: Laser Scanning Rangefinder. The left shows an outdoor LiDAR detection range of 30 m for intelligent robots (HOKUYO UST-30LX). The right shows LiDAR sensor mapping in a wide detection angle of 270°. Adapted from (HOKUYO, 2024).

In addition to the time-of-flight techniques in LiDAR systems, laser triangulation and phase-shift techniques are also used. Triangulation estimates distances from the angle of reflected light, offering a compact and cost-effective solution for indoor applications, such as RP Lidar A1 from Slamtec. Phase-shift measurement, in which distance is derived from the phase shift between the emitted and received laser signals, has been widely used in autonomous mobile robots, such as the Hokuyo UTM-30LX and SICK LMS-200.

2.1.2.4 RF sensor

Radiofrequency (RF) sensors in mobile robotics are used to determine positioning through triangulation or fingerprint techniques based on RSSI or RFID (Radio Frequency Identification) technologies. As Malyavej and Udomthanatheera (2014) described, RSSI is a promising method to derive a distance from a reference. The localization based on RSSI has been a popular research topic for a few decades due to the wide use of wireless communication in mobile phones and WLAN (Wireless Local Area Network). Furthermore, many WLAN transmitters can be used in most indoor environments without additional cost. According to Zhang et al. (2010), in most RFID positioning problems, the location information of RFID tags is interesting, but some applications involve the localization of RFID readers. An RFID positioning system generally consists of location sensing and positioning processing components.

The indoor localization based on radio frequency signal is categorized into two main types of techniques (Torteeke and Chundi, 2014): triangulation and fingerprint. The triangulation uses geometric properties of triangles to estimate the location, for example, AoA (Angle-of-Arrival), RSS (Received Signal Strength), ToA (Time-of-Arrival), RToF (Round-trip Time-of-Flight), and TDoA (Time Difference-of-Arrival), which the triangulation techniques researches were presented in the references (Torteeke and Chundi, 2014; Liu and Darabi, 2007; Vossiek et al., 2003). The fingerprint uses mapping RSS values (APs) and depends on sample location coordinates to generate a fingerprint map (Magrin and Todt, 2016; Yang and Zhang, 2014).

2.1.2.5 Digital Compass

Digital compasses are position sensors that measure the direction of a magnetic field. Digital sensors can operate according to the Hall effect or magnetoresistive effect.

The Hall effect describes the behavior of the electrical potential in a semiconductor when present in a magnetic field. It represents voltage according to the direction of the magnetic field and indicates the orientation of the magnetic field through the direction of the electrical potential. When a constant current is applied over the entire length of a semiconductor, there is a voltage difference across the width of the semiconductor based on the relative orientation of the semiconductor to the magnetic flux lines. Furthermore, the voltage potential sign identifies the magnetic field's direction. Thus, a single semiconductor provides measurement of flux and direction of the magnetic field (Siegwart et al., 2011). According to Siegwart et al. (2011), Hall effect compasses do not present good resolution, and internal sources of error include nonlinearity in the basic sensors and systematic adjustments in the circuit.

Sensors with magnetoresistive effect have the property that a magnetic material carrying current varies its resistance according to the application of an external magnetic field in a specific orientation. Figure 2.7 shows a strip of ferromagnetic material, called *permalloy* (19% Fe, 81% Ni) (Stork, 2000). During the deposition of the *permalloy* strip, a strong external magnetic field is applied parallel to the strip's axis. This defines a magnetization direction within the strip. Without an external magnetic field, magnetization always points in the x direction (current flow), and the resistance depends on the angle α between the current direction and the magnetization direction. In practice, the compass sensor is constructed as a Wheatstone bridge, which consists of four strips of magnetoresistive material, so the variation in resistance due to the magnetic field is converted linearly to a differential variation in the voltage output. Furthermore, the temperature coefficients of the four resistors are mutually compensated. According to Siegwart et al. (2011), regardless of the type of digital compass, Hall effect or magnetoresistive effect, the most significant disadvantage in using the Earth's magnetic field for applications in mobile robots involves external disturbances in the magnetic field by ferrous metals and magnetic objects.

However, the digital compass can provide helpful orientation information, even in the presence of steel structures.

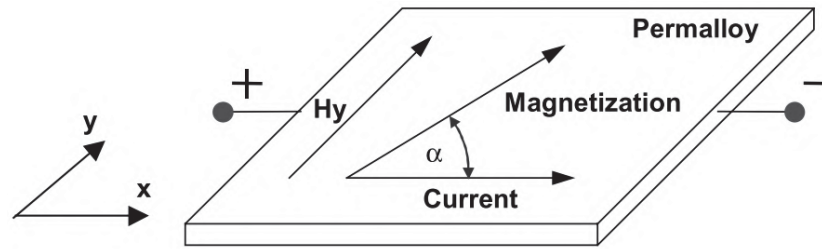


Figura 2.7: The magnetoresistive effect in *permalloy*. This is assumed to be the x-direction and the current flow direction. A magnetoresistive sensor now relies on two basic effects: The strip resistance depends on the angle α between the direction of the current and the direction of the magnetization; the direction of magnetization and, therefore, α can be influenced by an external magnetic field H_y , where H_y is parallel to the strip plane and perpendicular to the preferred direction. When no external magnetic field is present, the *permalloy* has an internal magnetization vector parallel to the preferred direction (Stork, 2000).

2.1.2.6 Ultrasound

Ultrasound *time-of-flight* is the most common technique for measuring distance in mobile robots in indoor environments. It is popular due to its low cost, lack of sensitivity to electromagnetic interference or smoke, and simple method of measuring distance (Siegwart et al., 2011; Murphy, 2000; Everett, 1995; Leonard and Durrant-Whyte, 1992). The ultrasound sensors are helpful in mobile robotics for measuring the distance, position, and location of objects, and avoiding collisions.

Figure 2.8 shows the low-cost ultrasound sensor HC-SR04, generally used in mobile robotics projects, and its main features are a measurement range between 2 cm and 400 cm with accuracy up to 3 mm and 15° measuring angle (ElecFreaks, 2004).

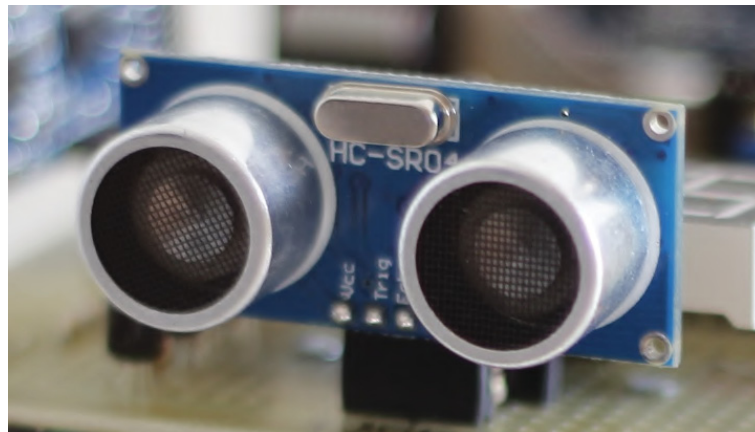


Figura 2.8: Ultrasound HC-SR04. The modules include ultrasonic transmitters, receivers, and control circuits. The HC-SR04 module provides a 2 cm to 400 cm measurement function, and the ranging accuracy can reach 3 mm (ElecFreaks, 2004)

The sonar opening angle, which forms a cone measurement, as shown in Figure 2.9, shows the typical intensity distribution of an ultrasound sensor. The shaded area covers the operating zone. Some disadvantages limit the use of ultrasound sensors, the most relevant of which is the possibility of *specular reflections* from the environment. The ultrasound sensor is

subject to *specular reflection*, producing an erroneous distance measurement at specific angles, even on a smooth wall. During robot movement, angles measured relative to a moving reference occur at random intervals (Siegwart et al., 2011). *Specular reflection*, Figure 2.10, occurs when the ultrasound transmitter signal cannot return directly to the receiver. The ultrasound sensors present a high conflict of false readings in distance measurement caused by *specular reflection*. False readings due to *specular reflections* can be treated as uncertainties from a probabilistic point of view. The ultrasound signal's incidence angle and the material's surface are decisive for *specular reflection*. When the incidence θ angle reaches a critical angle of *specular reflection*, the ultrasound fails to receive the correct value of the object's distance, i.e., the reading of a distance value is much greater than the actual distance value (Zou Yi et al., 2000; Drumheller, 1987).

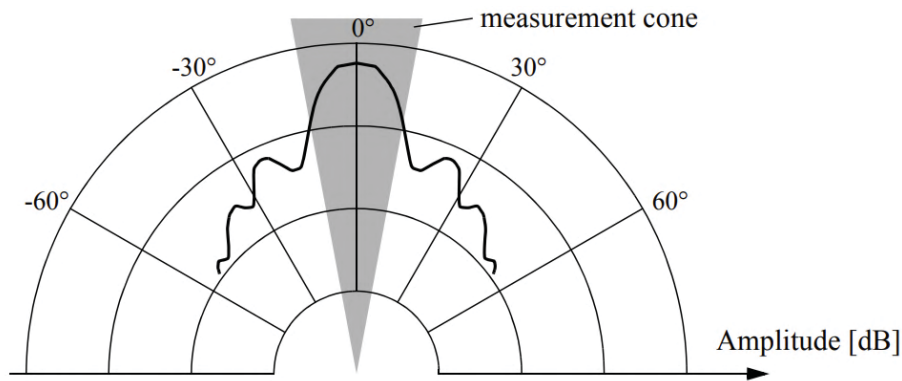


Figura 2.9: Typical intensity distribution of an ultrasound sensor. The figure shows the opening angle for the sound beam to obtain precise directional information about objects encountered. Ultrasound opening angle is a significant limitation since sound propagates in a conelike manner (Siegwart et al., 2011).

Figure 2.10 shows other relevant problems in ultrasound sensors: *crosstalk* and *foreshortening*. When carrying out measurements with different ultrasounds, interference known as *crosstalk* can occur - one ultrasound receives the signal sent by another close to it. *Crosstalk* interference can be resolved by changing the angle between sensors and the sensor reading sequence so that only the sensor that triggered the source signal receives the return signal. One of the characteristics of ultrasounds is that they have a cone-shaped field of view, generally at an angle of 30°. If the surface is not perpendicular to the sensor, one side of the cone reaches the object first and returns part of the wave as an echo, generating false readings. This problem is called *foreshortening* and occurs in measurements more significant than the measuring range (usually greater than 400 cm). This problem has no solution, so it is a feature of ultrasound measurements (Murphy, 2000).

2.1.2.7 Vision

According to Tzafestas (2014), mobile robot vision is the ability of a robot to see and recognize objects via the collection of an image of the light reflected by these objects and then interpret and process this image.

Several technologies involve computer vision, whether in supervising industrial processes or the perception of autonomous mobile robots. The main datasets used in mobile robotics, such as KITTI (Geiger et al., 2013), Oxford RobotCar (Maddern et al., 2017), and the University of Michigan NCLT (Carlevaris-Bianco et al., 2016), use cameras from Teledyne FLIR, a pioneer in developing thermal imaging solutions and infrared cameras. The main types of camera systems and commercial models used in mobile robotics are: monocular imaging - FLIR Grasshopper3 (Maddern et al., 2017), stereo imaging - FLIR Bumblebee® XB3 (Maddern et al., 2017), spherical

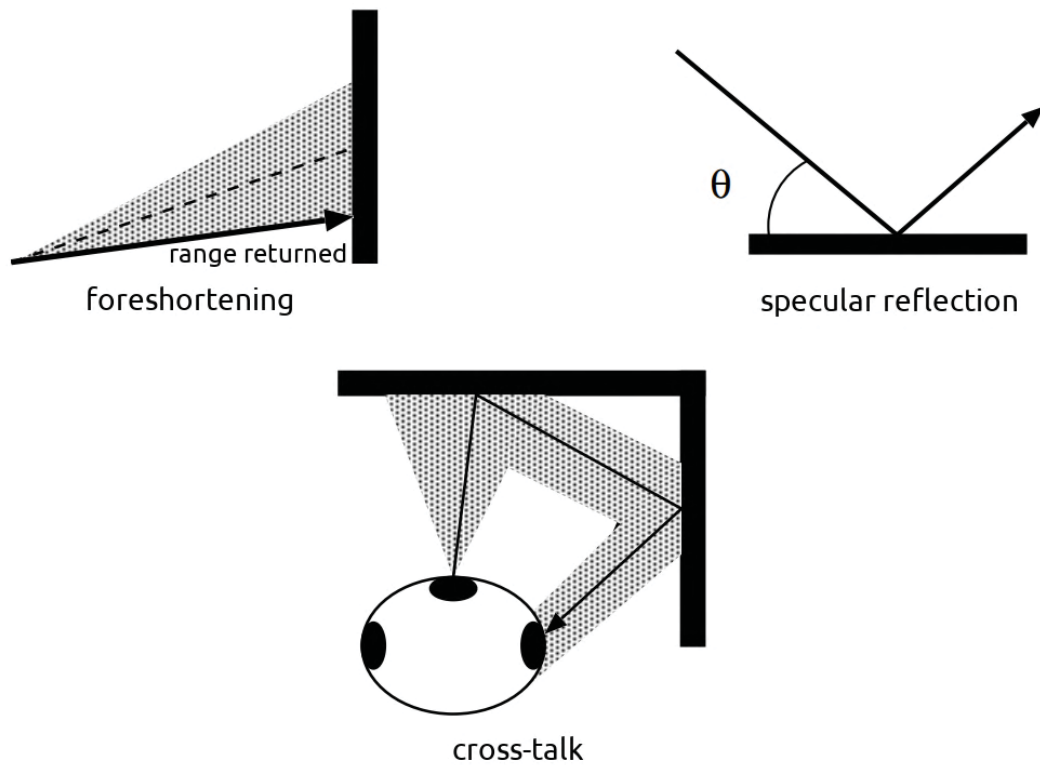


Figura 2.10: Three problems with ultrasound range readings: *foreshortening*, *specular reflection*, and *cross-talk*. *Specular reflection* - when the ultrasound transmitter signal cannot return directly to the receiver. *Crosstalk* - one ultrasound receives the signal sent by another close to it. *Foreshortening* - occurs in measurements more significant than the measuring range. Adapted from (Murphy, 2000).

imaging - FLIR Ladybug3 (Carlevaris-Bianco et al., 2016), depth imaging - Microsoft Kinect (Sturm et al., 2012), and point grey vision - FLIR Flea2 (Geiger et al., 2013).

Each camera sensor format reads the signal off given pixels differently, which can be a rolling shutter or a global shutter. The rolling shutter captures still images, recording sequentially and line by line. The sensor is simpler and has lower cost. The global shutter is used to capture moving objects, which is a global capture of everything simultaneously; each sensor pixel begins and ends the exposure simultaneously. Thus, a large amount of memory is needed, the sensor is relatively complex, and the price is expensive. Alkhawaja et al. (2019) presents the main visual and depth sensors used for localization systems and compares the specifications and cost of each device.

2.2 SENSOR FUSION

In mobile robotics, the field has many sensors for the perception of the indoor and outdoor environment, but each sensor has different features related to the perception of the environment. As Tzafestas (2014) described, sensor fusion is the process of merging data from multiple sensors to reduce the amount of uncertainty that may be involved in an autonomous mobile robot. Many types of sensors are currently available for multi-sensor fusion and embedded in an autonomous mobile robot: ultrasound, digital compass, accelerometer, gyroscope, laser scanner, and vision.

According to Khaleghi et al. (2013), sensor fusion has several issues that make data fusion challenging, such as data imperfection, outliers, conflicting data, data association, operational timing, and data dimensionality. Many of these problems have been identified because no single

sensor fusion algorithm can solve these issues. The variety of sensor fusion methods in the literature focuses on only a subset of sensors.

The sensor fusion is presented in (Luo and Kay, 1989; Zhao and Luo, 2008), considering that many autonomous mobile robot platforms are assembled with the goal of studying applications of multi-sensors and sensor fusion techniques. Multi-sensor fusion aims to achieve better system operation using the collective information from all sensors. Nowadays, the main details on fusion methods in the mobile robotics field include weighted average, Kalman filtering, particle filters, Bayes estimation, Dempster-Shafer method, Rough-set, neural networks, fuzzy logic, behavior-based algorithms, rule-based techniques, and hybrid fusion approaches as described in (Chandrasekaran et al., 2017; Zhao and Luo, 2008; Kam et al., 1997).

The literature suggests several configurations of multi-sensors. According to Durrant-Whyte (1988), a single source of sensory information can only provide partial information about an environment. The diverse information from many different sensors can be used to overcome the limitations inherent in using single sensors. The dynamic use of information between different sensors may occur in many ways and can be divided into three main categories: *competitive*, provide equivalent information about the environment, a typical sensor redundancy (For example, a configuration with three ultrasound sensors can tolerate the failure of one); *complementary*, merged information to form a more complex perception of the environment (For example, a set of ultrasound sensor covering non-overlapping measuring angle); *cooperative*, sensors work together to derive information that neither sensor alone could provide (For example, using two cameras in stereo for three-dimensional vision) (Durrant-Whyte, 1988; Brooks and Iyengar, 1997).

The data fusion may be arranged in three different architectures, such as *centralized*, the sensor fusion is treated as a central processor that collects all information from the various sensors, and a single fusion has the best performance if all the sensors are accurately, but minor errors in the sensors may cause substantial reduction in the performance of the fusion; *decentralized*, the sensor measurements, is fused locally using a set of local fusion rather than by using a single central fusion, but suffers from effects of redundant information; and *hierarchical*, as a hybrid architecture in which we mix the centralized and decentralized systems. In the hierarchical architecture, there are often hierarchical levels where the top level contains a single centralized fusion, and the last level is a decentralized fusion (Mitchell, 2007).

2.3 HIERARCHICAL SENSOR FUSION

According to Mitchell (2007), the hierarchical fusion consists of three levels of sensor management: sensor control, sensor scheduling, and resource planning. *Sensor control*, is the lowest hierarchical level. Focus on the individual control of each sensor, including the specification of its parameters, optimizing the sensor's performance, and data fusion given the present sensor measurements; *sensor scheduling* is the middle hierarchical level. Focus on the actual tasks performed by the sensors and on the different sensors. Prioritize the tasks that need to be performed and determine when a sensor should be activated; *resource planning* is the highest hierarchical level. Focus on tasks such as the placement of the sensors or the optimal mixture of the sensors required for a given task.

The main goal of hierarchical sensor fusion is to select a better set of sensors for each level of localization. To demonstrate related publications about multi-sensor fusion, follow the works of the authors:

Hierarchical sensor fusion method with fingerprint kNN and fuzzy features weighting (Magrin and Todt, 2016) was proposed to determine the location of an autonomous mobile robot. The sensor fusion method proposed is based on the fingerprint technique, feature weighting

with a fuzzy inference system, and a kNN matching algorithm. Mobile robots using low-cost sensors, such as eight ultrasounds with data invariant orientation from a digital compass, and more than fifteen available access points in the environment. The hierarchical sensor fusion allows a location in two levels: the first level divides the database and locates the robot's position in a larger space, and the second level determines its position and orientation (pose) in a smaller space in the map.

Simulation of a Mobile Robot Localization (Magrin and Todt, 2019b) has proposed a hierarchical sensor fusion method applied in a virtual robot experimentation platform (V-REP/CoppeliaSim) for indoor mobile robot localization. The proposal was to simulate a robot, Figure 2.11, equipped with low-cost sensors such as eight ultrasound sensors and a digital compass, using the received signal strength (RSS) from wireless networks available in the environment. The HSF method based on fingerprint kNN can determine the robot localization in different grid sizes. The development of this work with the ROS framework allowed segmentation into nodes for each application, such as teleoperation, perception, mapping, navigation, and localization.

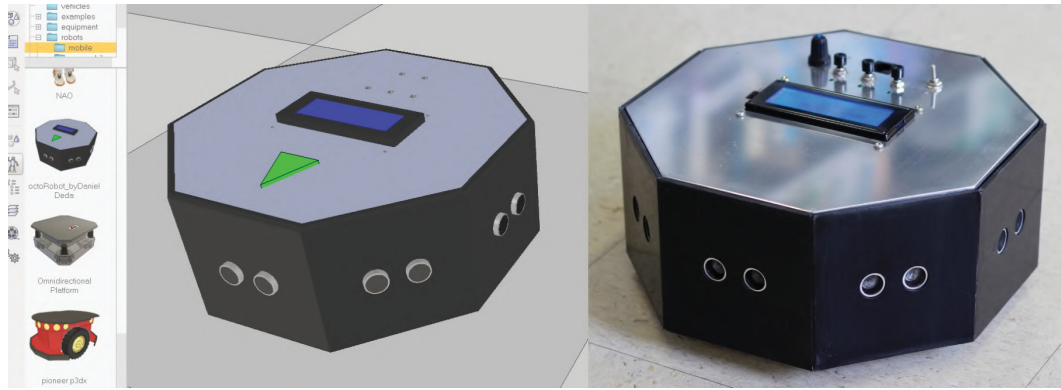


Figura 2.11: Mobile robot model in virtual platform V-REP/CoppeliaSim and physical OCTO robot platform. Simulating a two-wheeled mobile robot with proximity sensors (ultrasound), yaw orientation taken from the robot pose in the scene (digital compass), and included objects for representing the access points (APs) in the scene, for simulated received signal strength (RSS) from a wireless network.

Multi-sensor Fusion Method based on Artificial Neural Network (Magrin and Todt, 2019a) has proposed supervised learning, multilayer perceptron (MLP), and backpropagation technique to train the network in a hierarchical fusion step and determine the robot localization in a map. To validate this work, a comparison between the HSF methods, artificial intelligence, and the matching algorithm, using the same training and testing UFPR-RSFM Dataset (Magrin and Todt, 2019c).

According to Murphy (2000), the Hierarchical Paradigm is the oldest method of organizing intelligence in robotics. The Hierarchical Paradigm is represented by the relationship between the method *sense*, *plan*, and *act*, as shown in Figure 2.12. The Hierarchical Paradigm is sequential and orderly. The robot *senses* the world (perception), *plans* all the directives needed to reach the goal (planning), and *acts* to carry out the directives (controller). Figure 2.13 shows the reference control scheme (*see-think-act* cycle), depicted by Siegwart et al. (2011), for mobile robot systems. The robot senses its surrounding environment (perception), thinks about how to achieve its goal given the surrounding environment (path planning), and then acts (motion control).

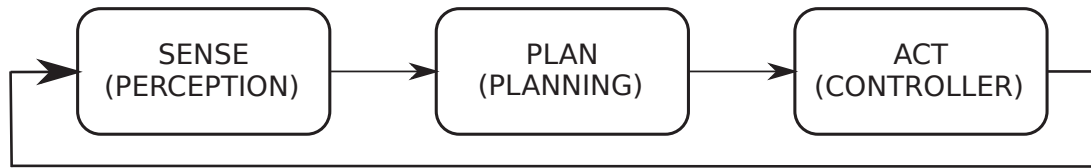


Figura 2.12: The hierarchical paradigm method. In the cycle, the autonomous mobile robot perceives the environment (sense), perform (plan), and controll the motion (act). Adapted from (Murphy, 2000).

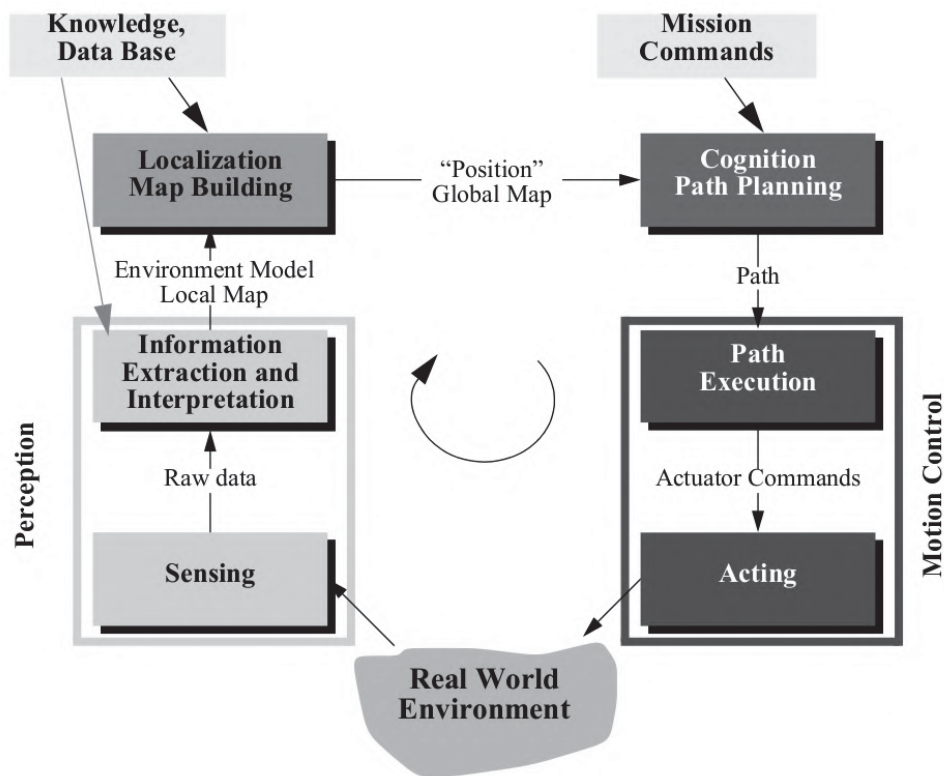


Figura 2.13: Mobile robots *see-think-act* cycle. The reference control scheme for mobile robot systems identifies much knowledge associated with mobile robotics, such as perception, motion control, localization, and path planning (Siegwart et al., 2011).

2.4 THE ARCHITECTURE OF THE MIND

In their works in psychology, the authors Newell (1994); Anderson et al. (2004) use the terms Architecture of the Mind and Theory of the Mind to evolve the perception-action system or, as represented in psychology, the perceptual-motor system.

Biology and medicine have found better answers to the fact that the body systems are specialized in their functions. According to Anderson et al. (2004), the specialist who studies the lung has a basic understanding of how their specialty relates to the specialist who studies the heart. The same idea, the computer scientist or engineer studies the decision-making and control of a robot. The different specialty areas, such as mechanics, electronics, and programming, are related. A basic understanding is essential for developing a mobile robot's perception, planning, and controller system.

According to Anderson (1995), one distinguishing feature of humans from other creatures is their ability to acquire complex skills, such as mathematics, language, chess, computer programming, and sculpture. Anderson (1995) presented a cognitive architecture, Adaptive Control of Thought (ACT), based on the unified theory of mind (Newell, 1994). This

cognitive theory aims to explain how human cognition works and the structures and processes of human memory, thinking, problem-solving, and language. ACT is a system that works on memory and perceptual-motor modules through buffers.

An ACT production system consists of three memories: working, declarative, and production. Working memory contains the information the system can access, consisting of the information retrieved from declarative memory deposited by encoding processes and the action of productions (Anderson, 1995).

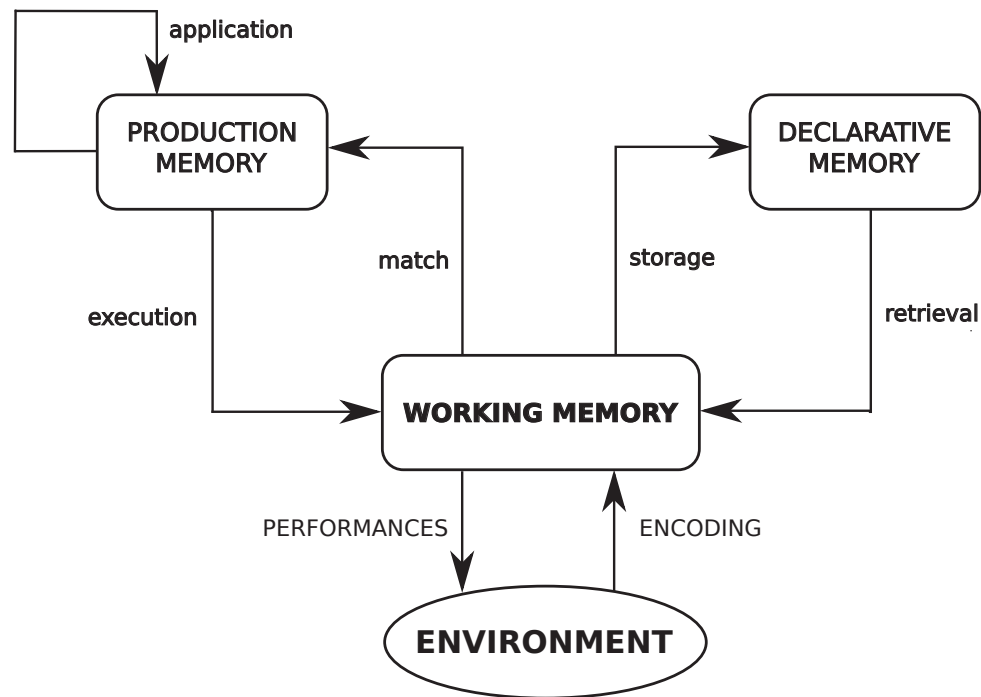


Figura 2.14: A general framework for the ACT production system, identifying the components and their interlinking processes with working memory. Declarative memory interlinks storage and retrieval, and production memory interlinks match and execution. Production memory feedback is the history of the application of existing productions. Adapted from (Anderson, 1995).

Figure 2.14 shows the basic process in the theory of cognitive architecture. The working memory involves the environment between *encoding*, processes deposit information about the environment into working memory, and *performance*, processes convert commands in working memory into behavior. The process involving memory and the environment is not the central theory of ACT. A general framework for the ACT production system identifies the components and their interlinking with the working memory, such as:

- *Storage* is a process that can create permanent records in declarative memory of the contents of working memory and increase the strength of existing records in declarative memory.
- *Retrieval* is a process that retrieves information from declarative memory.
- *Match* is a process of data in working memory that corresponds with the conditions of productions.
- *Execution* is a process that deposits matched productions' actions into working memory.
- *Application* is the new production learned from studying the history of the application of existing productions.

Working memory, knowledge to which we have access at any moment, can be identified with the active portion of ACT memory. This analysis consists of temporary knowledge structures currently being attended to and the active parts of long-term memory (Anderson, 1995). According to Anderson (1995), the cognitive system has evolved many different representational types, each intended to facilitate a specific type of computation. An analogy with computer science is that two programs can occupy different areas of computer memory, and two different cognitive abilities might lie in two separate regions of the brain. For instance, multiple sensory systems have evolved to perceive the same world. Human vision extracts different information but with overlapping signals, and cognitive processing optimizes these overlapping signals.

Cognitive psychology has two modes of cognitive processing: automatic and controlled. Automatic, less capacity-limited, possibly parallel, invoked directly by stimulus input, and control with severe capacity limitations, is perhaps serial, and is invoked in response to internal goals (Anderson, 1995). The current version of Adaptive Control of Thought – Rational (ACT-R) is based on the principle of rationality of the human mind (Anderson et al., 2004). Rationality principle is defined by Newell (1994): *“If an agent knows that one of its actions will lead to one of its goals, then the agent will select that action”*. ACT-R evolved multiple modules and explained how these modules are integrated to produce coherent cognition. In control of cognition, the hierarchical processing *bottom-up* processing starts with the data and tries to work up to the high level, and *top-down* processing tries to fit high-level structures to the data.

How do we describe the basic structure of cognitive systems? According to Newell (1994), our ability to explain human cognition in one way rather than another rests ultimately on human beings’ physical and biological nature. The work of Newell (1994) represents the foundations of cognitive science and analyzes the main fundamentals of the theories. The theories cover behaving systems, knowledge, representation, computational systems, symbols and symbol systems, architecture, intelligence, search and problem space, and preparations and deliberation.

Behaving system is characterized by multiple response functions on its environment. The mind can be seen to provide response functions, and the control system guides the behaving organism in its complex interactions with the dynamic real world.

Knowledge level can be used to select actions to reach the goals. Figure 2.15 shows the agents with knowledge and goals. Newell (1994) says: *“The behavior of an existing system can be calculated if you know the system’s goals and what the system knows about the environment”*. The knowledge levels represent specifications for the systems, and each level is helpful to serve as a specification for the lower level.

Representation is a system that accepts representation laws. Figure 2.16 shows a hierarchy of computer systems, for example, a knowledge-level system that exists in the physical world, a program system, a register-transfer system, a logic system, an electrical circuit, and an electronic device. Knowledge abstracts from representation, yet knowledge must be represented in some fashion to be used. The standard way of representing knowledge is with logic. The world of the organism can be in many different states. Even if correlations exist between descriptors, and the organism attends only selectively to the environment, the resulting possibilities that must be encoded in any situation remain combinatorial. According to Newell (1994), selective attention restricts variety. It simultaneously contributes to combinatorial variability by allowing the system to be focused on one place or another, independently of the state of the world, so that total variety becomes a possibility for attention versus options for the environment.

Computational system is a machine that obtains multiple functions on some inputs and outputs by using other inputs to specify functions. The general notion of a system as a region of the physical world that can be described by a set of interacting variables.

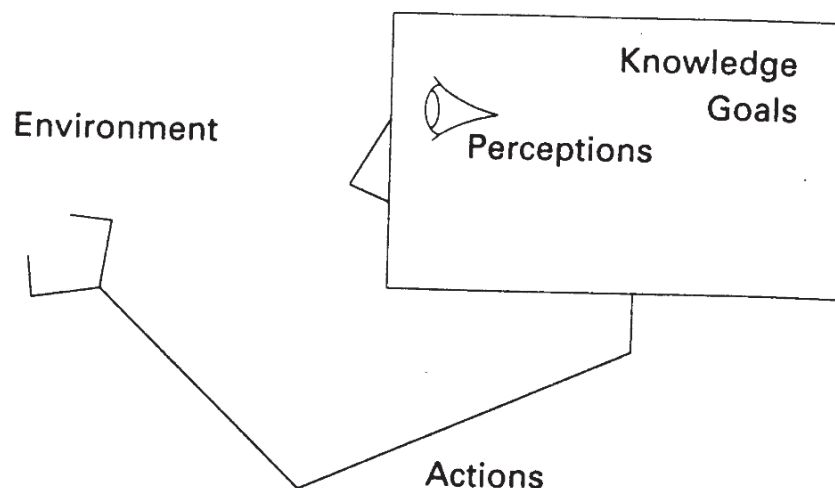


Figura 2.15: The knowledge-level system. The agent, perception of the environment, and action consider the goals and knowledge. The system represents the concept of hierarchical paradigm method used in the autonomous mobile robot, sense-plan-act cycle (Figure 2.12) and see-think-act (Figure 2.13) (Newell, 1994).

Symbols represent parts of structures to be processed. Anything that represents exhibits a symbolic function. In cognitive science, Symbol systems are a form of universal computational system. A symbol system is sufficient for producing all computable functions. The variety of knowledge and goals to be represented is large enough. Newell (1994) presents the research questions. How do the symbols in symbol systems represent something external to the symbol system? How do symbols acquire their semantic function of referring to things in the outer world?

Architecture is the fixed structure that realizes a symbol system. The unified theories of cognition are formulated as architectures. At the top of the computer systems hierarchy is the knowledge level, below that, the symbol level, and then the register-transfer level or the architecture. According to Newell (1994), computer science has built up the notions of virtual machine, firmware, and dynamic configurability as appropriate ways to describe such structures. The computer systems hierarchy is an essential invariant structural characteristic.

Intelligence is another central concept of cognitive science. The natural definition of intelligence, as Newell (1994), is a system that is intelligent to the degree that it approximates a knowledge-level system. Intelligence is the ability to bring to bear all the knowledge that one has in the service of one's goals; that is, intelligence is relative to goals and expertise.

Problem space search in the problem space to solve a task. The search represents the knowledge search - search in memory to guide problem search and problem space - a set of states and a set of operators from state to state. Search is fundamental for intelligent behavior. The discussions of psychological processes involved in thinking, reasoning, and intelligence before the advent of Artificial Intelligence in the mid-1950s show no particular emphasis on search. The intelligent behavior is deeply involved with determining what spaces are searched and controlling search in those spaces. The character of any specific search depends on many things, such as the memory available or the types of search steps that are possible.

Preparation vs. Deliberation trade-off is shown in Figure 2.17, where the axes represent different forms of knowledge acquisition. The vertical axis represents immediate knowledge, which is encoded in the system's memory (Preparation). Conversely, the horizontal axis represents search knowledge, obtained through Deliberation by considering multiple situations via searching. The trade-off highlights a state where an operator has been selected, but

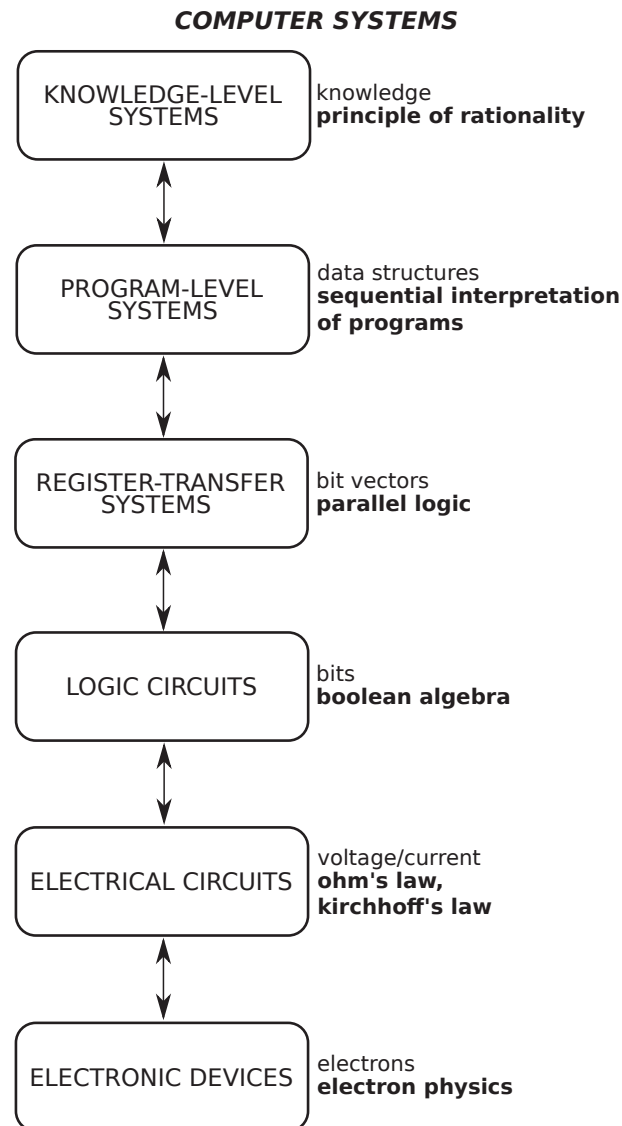


Figura 2.16: The hierarchy describes a computer system as a system of electronic devices with electron physics laws, electrical circuits with Ohm's and Kirchhoff's laws, logic circuits with Boolean algebra, register-transfer systems with parallel logic, or programming systems with a sequential interpretation of programs in many knowledge levels. Adapted from (Newell, 1994)

immediate knowledge is insufficient to produce a new state, necessitating reliance on the search knowledge.

2.4.1 The total cognitive system

Newell (1994) presented the total cognitive system, Figure 2.18, the cognitive model is the principle of *see-think-act* cycle, used by Siegwart et al. (2011) in the reference control scheme for mobile robot system (Figure 2.13). In central cognition or *cognitive*, showed in Figure 2.18, the main characteristics are: problem spaces represent all tasks, solving architecture; the recognition memory where productions provide all long-term memory in search control, operators, and declarative knowledge; attribute-value representations serving as the medium for all information processing; a preference-based procedure is used for all decision cycle, for example, accept/reject and better/worst; goal stack direct all behavior and operators perform the

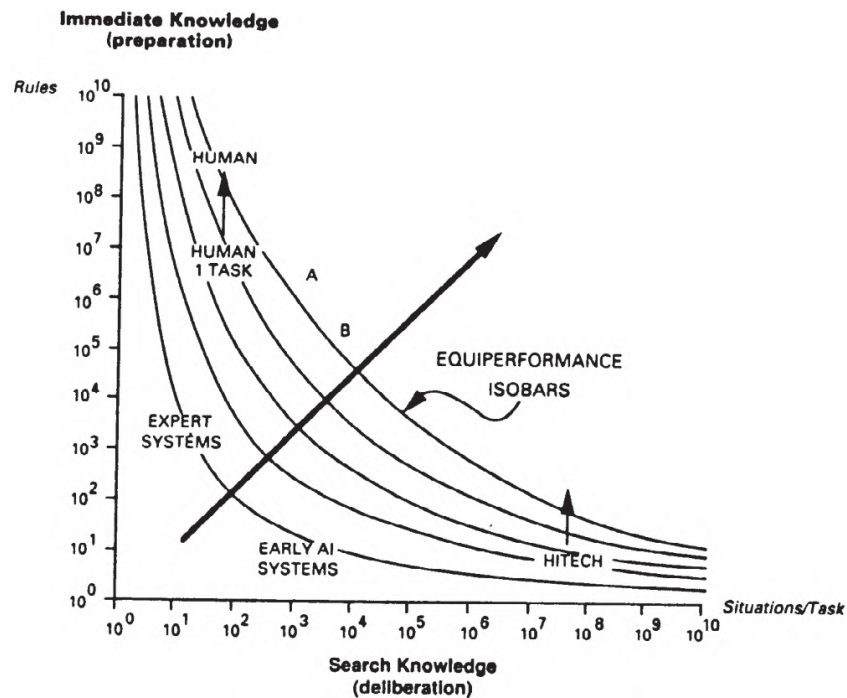


Figura 2.17: Preparation versus deliberation trade-off. The space trade-off distributes limited development resources in a one-dimensional manner. The curves represent the AI systems, expert systems, and human tasks. Humans have better immediate knowledge (preparation), and computers (HITECH) have better search knowledge (Newell, 1994).

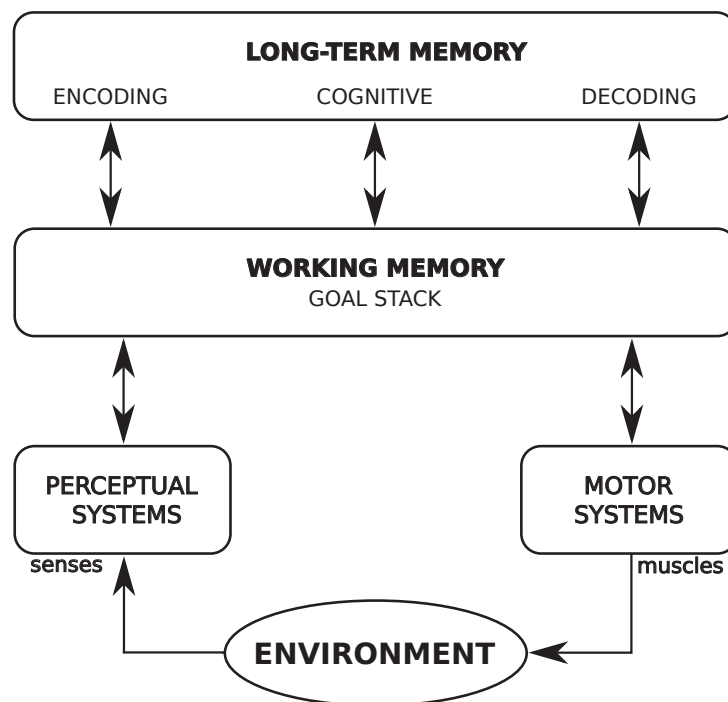


Figura 2.18: The total cognitive system: perception, cognition (long-term memory), and the motor system. The perceptual systems sense the environment and buffer it in working memory, encoding, parsing, and putting the information into the cognitive process. The decoding provides commands produced by the cognitive system, and working memory provides information available to internal components for use by the motor system. Adapted from (Newell, 1994)

function of deliberate goals; and continuous chunking processes that consolidate goal-results. This larger system includes perception and motor behavior.

According to Newell (1994), the mind is a controller of a dynamic system interacting with a dynamic environment. In the *perceptual system*, the process transduces the energies in the environment into signals for the system. The process that affects the environment in the *motor system*. *Cognitive* in the working memory becomes the communication device for all the total system components (buffer). Buffering on the perceptual side puts elements into some special memory. Another particular device could move them into the working memory. On the motor side, buffering makes information available to internal components. In the perceptual system, the encoding provides perceptual parsing, putting the elements into a form to be considered by *cognitive*, and the decoding provides the expansion and translation of the motor commands produced by the cognitive system into the form used by the motor system.

The attention operator is the active process and can be executed by central cognition. Given a search specification, the attention operator finds an element satisfying this specification from anywhere in working memory. The result of the attention operators is that it switches cognition to the place of the attended feature. According to Newell (1994), the attention operator is like a channel selector. Parts of the working memory into which the perception systems put their elements are organized as topographical fields, such as metric spaces with relations of at, near, far, and direction. The attention operator is an architectural mechanism with primitive capabilities corresponding to the attributes made available by perception. The attention operator is voluntary and can be chosen by deliberation, just as any operator, but attention can also interrupt cognition. The architecture makes the best preference for the attention operator selected at the next decision cycle.

According to Anderson et al. (2004), although human cognition is certainly embodied, its embodiment does not give human cognition its advantage over other species. Its advantage depends on its ability to achieve abstraction in content and control.

Meyer and Kieras (1997) developed a scheme for relating cognition to perception and action without dealing directly with real sensors or real effectors and without having to embed all the details of perception and motor control. They introduced an Executive-Process Interactive Control (EPIC), concurrent perceptual-motor and cognitive tasks, and a comprehensive theoretical framework for developing precise computational models and applying them to characterize human multiple-task performance under various conditions. Figure 2.19 shows the EPIC architecture, consisting of complementary memory stores and processing units that interact hierarchically. Memory stores have three functionally distinct memory stores: declarative long-term memory, production memory, and working memory. Processing units are visual, auditory, and tactile perceptual processors that receive inputs from simulated physical sensors. Each perceptual processor sends outputs to working memory, which a cognitive processor uses to perform various tasks. The cognitive processor selects symbolic responses and sends them to vocal and manual motor processors, which prepare and initiate movements by simulated physical effectors (Meyer and Kieras, 1997).

In the EPIC architecture (Figure 2.19), each component has symbolic representations, input/output transformations, and duration time to model human performance. Meyer and Kieras (1997) presents each type of component and its properties, such as:

- *Perceptual processors* operations are parallel and asynchronous; stimulus identities are sent to working memory; transmission times depend on modality, intensity, and discriminability.
- *Motor processors* respond to identities received as inputs; movement features are prepared for physical outputs; feature preparation is done serially with set time increments; advanced feature preparation is done for anticipated responses; movement initiation is

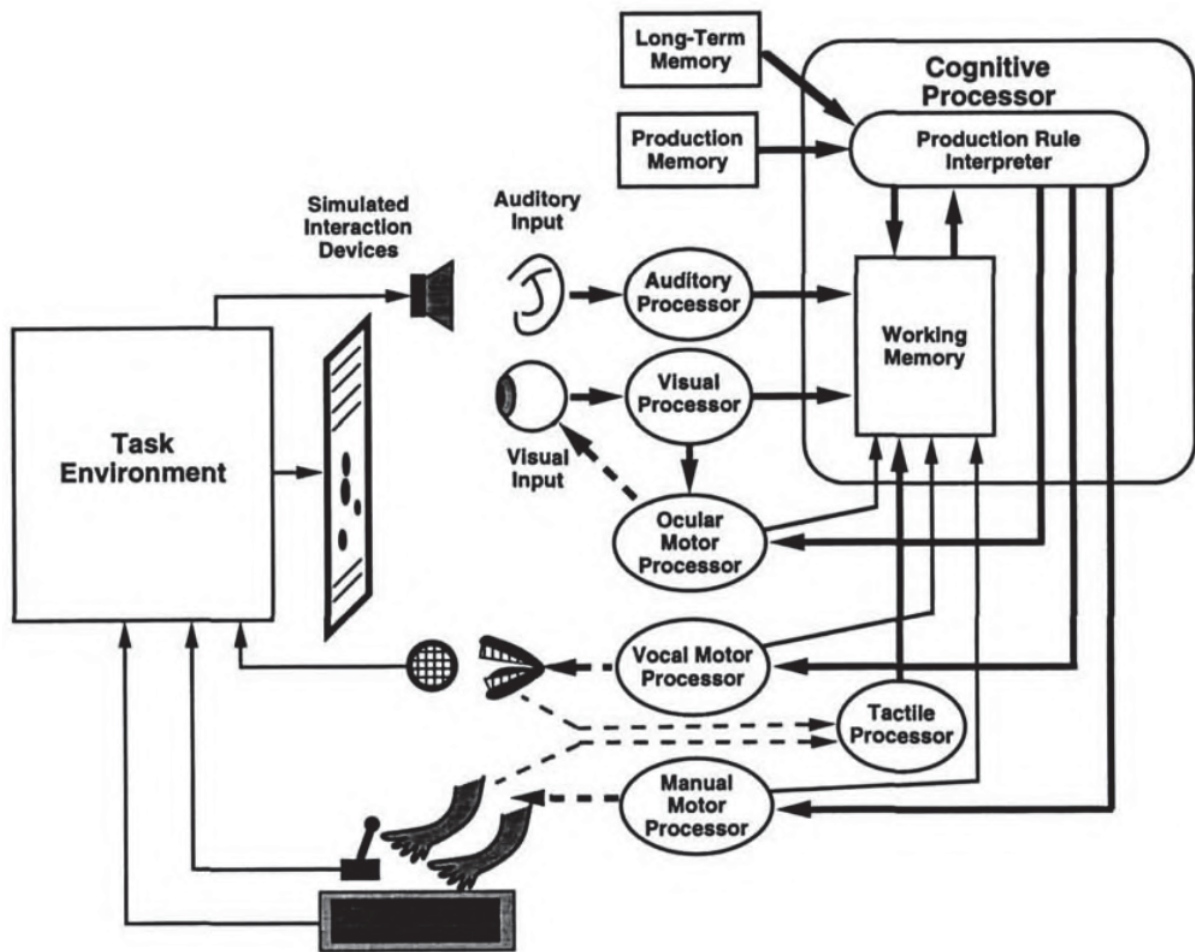


Figura 2.19: Overview of the information-processing components in the executive-process interactive control (EPIC) architecture. The figure shows the human cognition model, interaction with the task environment, input in perceptual processors (auditory, visual, and tactile), output in motor processors (vocal, manual, and ocular), computer in cognitive processor production rules, and working memory consists of goals (Meyer and Kieras, 1997).

done after feature preparation; efference copies of motoric representations are sent to working memory.

- *Cognitive processor* programmed with production rules; rules interpreted by parsimonious production system; conditions refer to goals, steps, and notes in working memory; steps in conditions govern flow of control; complex conflict-resolution criteria and spreading activation not used; actions regulate working memory and perceptual-motor processors; cyclic operation with set mean cycle duration; no limit on number of rules being tested and applied simultaneously.
- *Working memory* consists of goals, steps, and notes; contents used and managed by a cognitive processor; capacity and duration sufficient for performance in Psychological Refractory Period (PRP) procedure.

Response selection refers to a subsequent process that converts the stimulus code to an abstract symbolic code for a physical response based on some innate or previously learned stimulus-response associations. People have limited numbers of sensors (two eyes and two ears), and effectors (two hands and one mouth), and the constraints imposed by them are essential to understanding multiple-task performance (Meyer and Kieras, 1997).

Cognitive systems theory has been revised in psychology to support the next section on cognitive systems. We can better understand the human cognitive system and the agent based on dynamic cognitive systems. For the interested reader in psychological review, the Theories of Cognition and Theory of the Mind are presented in detail in the references (Newell, 1994; Anderson, 1995; Meyer and Kieras, 1997; Anderson et al., 2004).

2.5 COGNITIVE SYSTEMS

The cognitive systems, inspired by human cognition and cognitive dynamic systems (Fuster, 2005; Haykin, 2012), explore four cognitive processes: perception, memory, attention, and intelligence. The basic processes involved in the cyclic operation of the cognitive system (Figure 2.20) (Haykin, 2012; Fuster, 2005) are:

- *Perception* for a cognitive system to perceive the environment, the system must be equipped with an appropriate set of sensors to learn the environment, aimed at learning the underlying physical attributes that characterize the environment (environment scene analysis).
- *Memory* is dynamic in a cognitive system in that its contents continually change over time following changes in the environment. Memory contents are subject to time constraints, whereas knowledge is a memory of specific facts and relationships that exist between them and is timeless in its contents.
- *Attention* is a mechanism for prioritizing resource allocation (data) selectively in terms of practical importance. Protect the system from information overload problems by prioritizing how these computational resources are allocated.
- *Intelligence* is the ability of a cognitive system to continually adjust itself through an adaptive process by making the perception respond to new changes in the environment to create new forms of action and behavior.

Perception can lead to other cognitive processes or behavioral actions. Perception to action is translated through connections between sensory and motor structures. The perception-action cycle involves the cyclic operation of the cognitive systems (Fuster, 2005).

2.5.1 Perception-action cycle

According to (Fuster, 2005), the cognitive interactions of a human with the surrounding world are governed by the perception-action cycle (Figure 2.21). This interactive cycle is the extension to cortical processes of a basic principle of biology that characterizes the dynamic adaptation of an organism to its environment.

To interact adaptively and selectively with the environment, the cerebral cortex requires the functional integrity of subcortical sensory and motor systems (Fuster, 2005).

2.5.1.1 Sensory hierarchy

Sensory is an area dedicated to receiving and storing information from the senses, as well as the cognitive functions related to them. According to Fuster (2005), the sensory hierarchy observes the surrounding environment in three hierarchical levels inspired by the sensory cortex. The structure of the cortex around the diagram perception-action cycle shows *primary sensory cortex*, dedicated to the representation of an elementary sensory feature; *unimodal association*

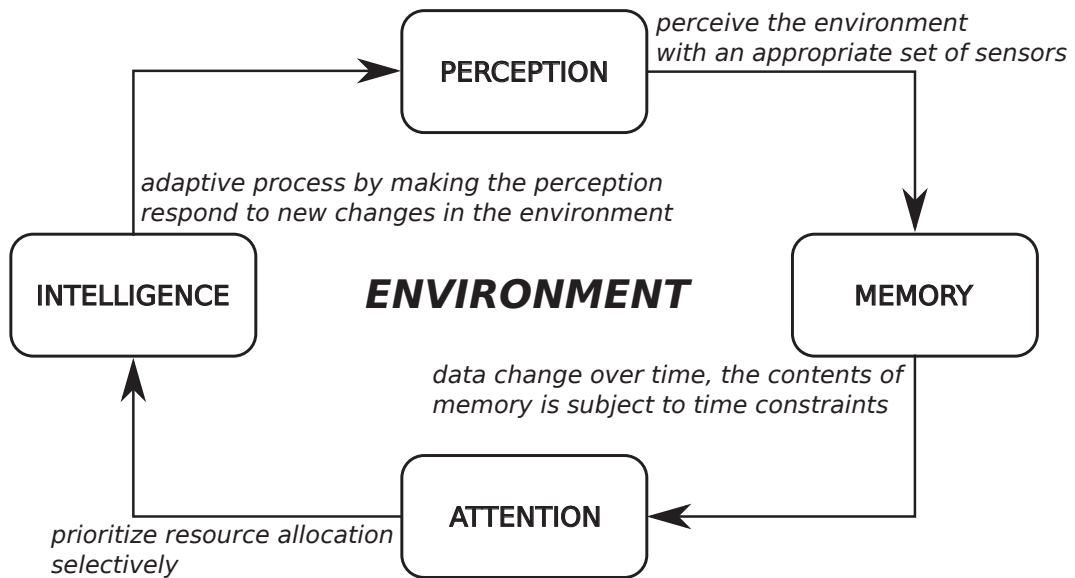


Figura 2.20: Cyclic operation of the cognitive system. The cyclic process involves the environment and connects the systems. Perception, the cognitive system perceives the environment; Memory, with the data changing over time, the contents of memory are subject to time constraints; Attention, as the computational resource is limited (memory), the attention system prioritizes resource allocation; Intelligence, the adaptive process continually adjusts the plan.

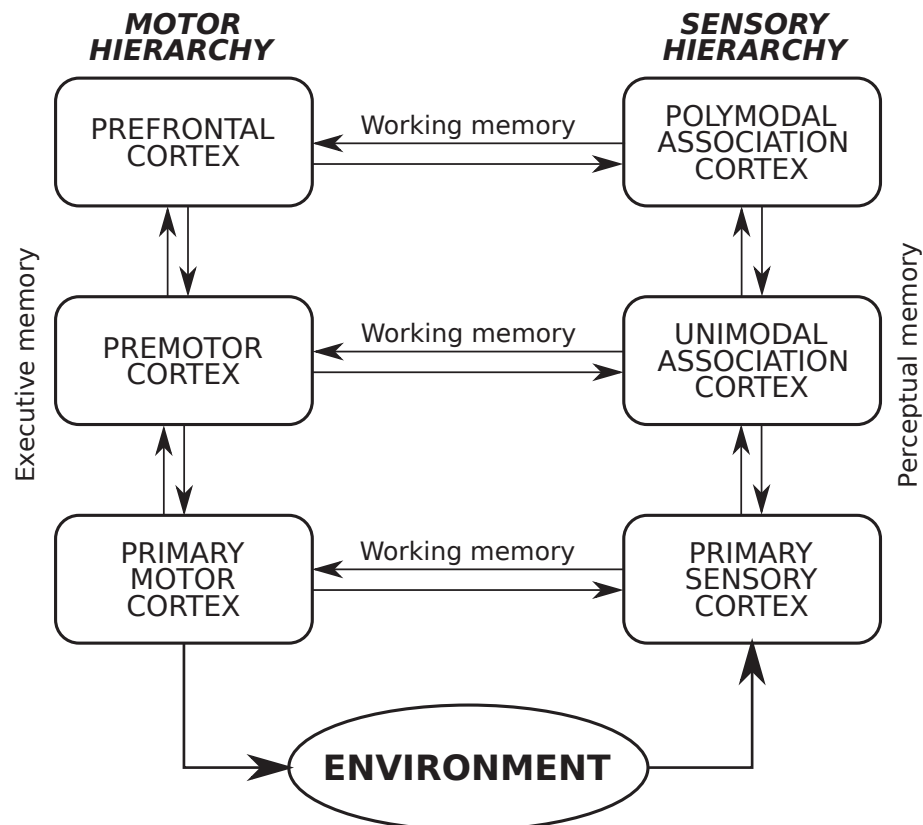


Figura 2.21: Perception-action cycle of a cognitive system. The sensory hierarchy observes the surrounding environment in three hierarchical levels: primary, unimodal, and polymodal. The motor hierarchy is an area dedicated to the system's actions, involving the cortex's primary, premotor, and prefrontal. Working memory represents a recent event for a pending action, and executive and perceptual memory include all knowledge acquired through a learning process. Adapted from (Haykin, 2012; Fuster, 2005).

cortex, a significant level in the hierarchy dedicated to the processing of complex information; *polymodal association cortex*, integration of information coming in through different senses. Perceive their self-motion and environment to locate and navigate a complex three-dimensional space.

2.5.1.2 Motor hierarchy

The motor is an area dedicated to the actions of the system and its related cognitive functions, inspired by the motor cortex (Fuster, 2005). Perception-action diagram showing *primary motor cortex*, generate neural impulses that control the execution of movement; *premotor cortex*, direct control of motor behavior, playing a role in planning movement and using rules to perform specific tasks; *prefrontal cortex*, control executive functions, such as planning and decision making.

2.5.2 Memory

Information contents in the memory of the cognitive system continually change over time following changes in the environment. The perception-action cycle of a cognitive system, shown in Figure 2.21, consists of perceptual memory, executive memory (action), and working memory. The links between hierarchies and stages retrieve old memories (database) about information in the past and update them with new information about the environment.

Memory is stored in both sectors: perceptual memory and executive memory. Working memory represents a recent event for a pending action (Fuster, 2005).

- *Perceptual memory* includes all knowledge acquired through sensory learning.
- *Executive memory* includes all knowledge acquired through actions learned.
- *Working memory* represents the temporary memory of an online event for solving problems in the perception-action cycle and feedback information, executing cognitive operations between the sensory and motor hierarchy.

One single perception of the environment can access memory at several hierarchical levels. That multilevel access to memory networks is why, when access to one level of memory is reinforced, access to another becomes facilitated. The reason why the indicating or organizing of memory items by category makes it easier to recall (Fuster, 2005).

2.5.3 Attention

Cognitive functions are sensory and motor interactions with the environment. It is a perceptual processing that is guided by selective attention (Fuster, 2005). Attention is a cognitive mechanism that prioritizes resource allocation and helps avoid information overload. In other words, the purpose of attention is to allocate selectively features to the system.

Based on the cognitive dynamic system (Haykin, 2012), the role of attention in cognition is that the system allocates the available resources, including prior knowledge, and prioritizes the allocation in order of importance using perceptual memory and executive memory. In addition, attention looks to the working memory for information on the consequences of actions taken by the system.

The essence of attention is the selective allocation of available resources necessary for the perception analysis. In cognitive sensor fusion, the attention represents sensor selection

according to environmental characteristics. Attention does not generate new sensory inputs. However, information can be better analyzed and subsequently brought in for analysis (Fuster, 2005). According to (Berto et al., 2020), in robotics, attention could be perceived as the capacity to filter information received according to parameters of interest and objectives, reducing the agent's search space. An example of attention in robotics is the pioneering work of Todt and Torras (2004), using attention in the visual field. The proposal reduces the amount of information processed to analyze a scene image to find reference landmarks for the outdoor localization of autonomous mobile robots.

2.5.4 Intelligence

In cognitive functions, intelligence can be defined as adjusting by reasoning to new changes, solving new problems, and creating new forms of action. From its beginning, the purpose of artificial intelligence was to construct models of Turing machines, computers, and robots that performed cognitive functions like those of the human brain (Fuster, 2005). The other cognitive processes (perception, memory, attention, and language) also contribute to intelligence.

2.6 ROBOT OPERATING SYSTEM (ROS)

Robot Operating System (ROS) is an open-source framework for writing robot software, and it simplifies the task of dealing with a wide variety of robotic platforms. Although ROS is not an operating system, it provides services designed for hardware abstraction, low-level control, high-end capabilities, and packages with tools for debugging, data visualization, and interface simulation. As Guimarães et al. (2016) described, ROS has a set of valuable resources so a robot can navigate through a medium. The robot is capable of planning and following a path while avoiding obstacles that appear on its path throughout the course. According to OpenRobotics (2021), ROS offers a standard software platform to developers across industries that carries them from research and prototyping to deployment and production. ROS is ready for use across various robotics applications, from indoor to outdoor, home to automotive, underwater to space, and consumer to industrial. Many ROS groups have grown to the point where they are large enough to be considered their standalone organization, such as micro-ROS, bridging the gap between microcontrollers and processors in robotic applications; OpenCV (Open Source Computer Vision Library), an open-source computer vision and machine learning software library; Open Robotics, which works with the global ROS community to create open software and hardware platforms for robotics including ROS, ROS 2, and the Gazebo simulator; ROS-Industrial, an open-source project that extends the advanced capabilities of ROS software to industrial relevant hardware and applications.

The main changes between ROS and ROS 2 are that ROS is only being tested on Ubuntu, targets Python 2, and uses a custom serialization format and transport protocol; ROS 2 is currently tested on Ubuntu and supported on Windows, requires at least Python version 3.5, and provides various Quality of Service policies that improve communication over different networks. This work started before ROS 2 was stable, so several packages were tested using ROS. During the development of this work, we judged the feasibility of changing to ROS 2, but the update would not benefit the development in question. We may encounter issues using some packages that are not updated for ROS 2.

ROS, used in this work, is represented by a service called *roscore* that provides connection information to *nodes* so that they can transmit messages among themselves through a common topic to *publish* and *subscribe* data. Figure 2.22 shows the NODE registration information to ROS MASTER, and TOPIC gives or takes a message from the node.

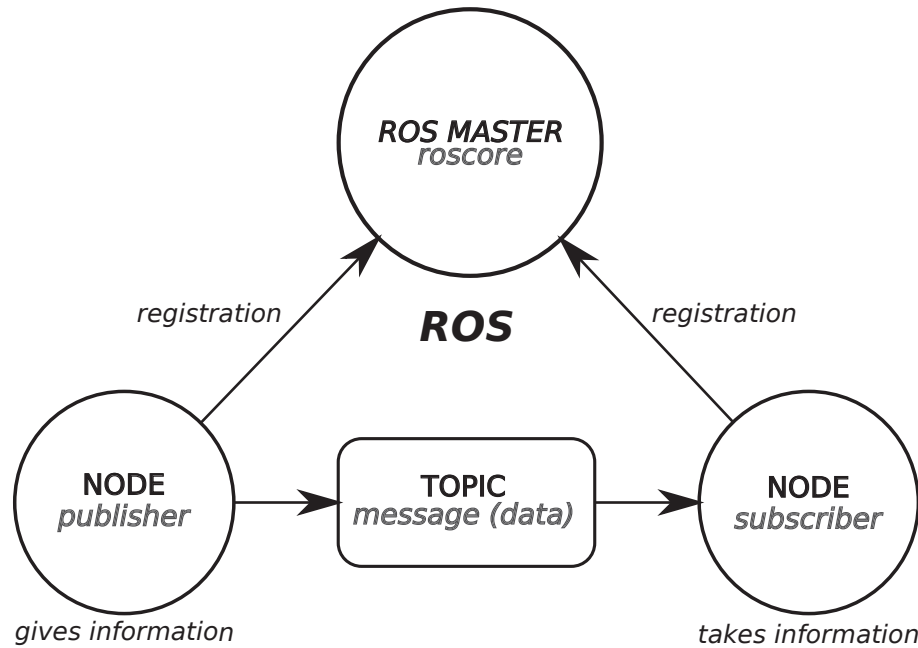


Figura 2.22: Understanding ROS Nodes. The ROS MASTER is a service for ROS (*roscore*) is the first thing you should run when using ROS. A node is an executable that uses ROS to communicate with other nodes. Nodes can *publish* messages to a topic and *subscribe* to a topic to receive messages. Messages are ROS data used when *subscribing* or *publishing* to a topic. Adapted from OpenRobotics (2021).

2.7 DIGITAL TWIN

According to Tao et al. (2019a), the relevance of the *digital twin* concept is increasingly emphasized by academia and industry, related to the *cyber-physical* systems, and is one of the challenges for smart manufacturing and Industry 4.0 for connecting the physical and virtual environment. In the study of Tao et al. (2019a), there are different understandings of *digital twin*; some researchers focus on simulation, others on a basic three-dimensional concept (physical, virtual, and connection) proposed by Grieves (2014). Tao and Zhang (2017) proposed to include five dimensions (physical part, virtual part, connection, data, and service). The *cyber-physical* fusion introduced by the *digital twin* concept provides the framework and systems for interaction between the real world and virtual environment. *Digital twin* provides the mobile robot researcher with flexibility in evaluating learning algorithms and reducing initial costs in the physical robot design so that the robotic platform can first be tested and evaluated in a virtual environment, to integrate later the *cyber-physical* system, resulting in the quality and reliability of the trained data and mobile robot evaluation.

Barricelli et al. (2019) described a *digital twin* as a part of a *cyber-physical* system in smart manufacturing and Industry 4.0, which can be a set of physical entities (devices, objects, equipment, humans) that interact with virtual cyberspace through a communication network. The modular architecture of virtual twins (simulation) in applying data fusion algorithms, big data analytics, and artificial intelligence algorithms evolves with its physical twin, allowing a fast system reconfiguration. It is essential to understand that the *digital twin* is not always fully autonomous and still requires a lot of human intervention, particularly in environments where they are used to test new features and modifications of physical assets or when exploited to provide answers such as diagnosis and treatments.

The concept of *digital twin* was revisited by the National Aeronautics and Space Administration (NASA), which defined the *digital twins* as “an integrated multi-physics, multi-

scale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its flying twin. It is ultra-realistic and may consider one or more important and interdependent vehicle systems” (Glaessgen and Stargel, 2012).

Pires et al. (2019) introduced the concept of *digital twin*, creating a virtual copy of the physical system, providing a connection between the real and virtual systems to collect, analyze, and simulate data in the virtual model to improve the performance of the real system, which impacts the simulation of the model. Figure 2.23 shows the case study of Pires et al. (2019), the applicability of the digital concept in the virtualization of a UR3 collaborative robot that used the V-REP simulation (CoppeliaSim) environment and the Modbus communication protocol, real and virtual robots moving at the same time. The architecture consists of the basic three-dimensional concept, a physical part (the UR3 robot), the virtual part (the model of the robot represented in the V-REP simulation software), and communication (custom Java application), where the communication between the real robot and the virtual model is on the V-REP simulation platform. The *digital twin* can speed up the commissioning process, such as design, testing, adjusting, verifying, and training. It is also responsible for performing robot optimization, reducing the risk of the project, and performing early fault detection (Pires et al., 2019). The simulation of a mobile robot proposed by (Magrin and Todt, 2019b) modeled a robot in a virtual scene and segmented the project in ROS nodes, supporting a sensor fusion method, simulation in CoppeliaSim, and solves the mobile robot localization problems in indoor environments using the concept of the modular architecture of virtual twin (simulation) for flexibility, portability, and scalability of the simulation model.

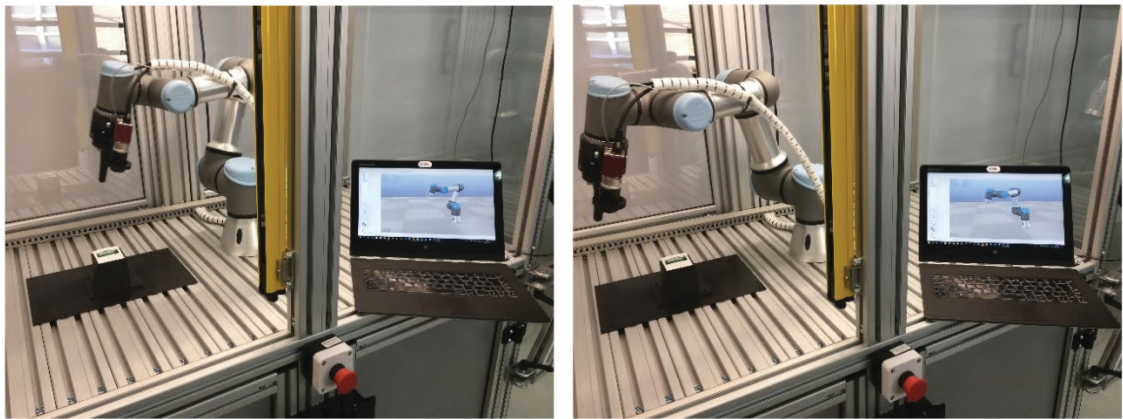


Figura 2.23: Shows a real and virtual robot moving simultaneously. The architecture consists of the basic three-dimensional concept, a physical part (the UR3 robot), the virtual part (the model of the UR3 robot), and communication (Modbus communication protocol in Java application) (Pires et al., 2019).

2.7.1 Digital Twin Mobile Robot

The traditional three-dimensional *digital twin* attends to the basics of creating a real mobile robot modeled in a virtual environment (physical, virtual, and connection). Furthermore, for creation as a tool for mobile robotics for this work, using the framework ROS, the five-dimensional model is more suitable. We used the concept with a full five-dimensional digital twin of Tao et al. (2019a) for proposing a **Digital Twin Mobile Robot** (DTMR) - presented at the work of Magrin et al. (2021). Figure 2.24 shows the concept of the DTMR, using a *physical* system (real mobile robot in a real environment), a *virtual* system (mobile robot model in a virtual scene), *service* (integration between the ROS nodes), *data* (exchange ROS message by

topics) and *connection* (data exchange, and running programming using the same standard data and information in both *cyber-physical* systems through the ROS master).

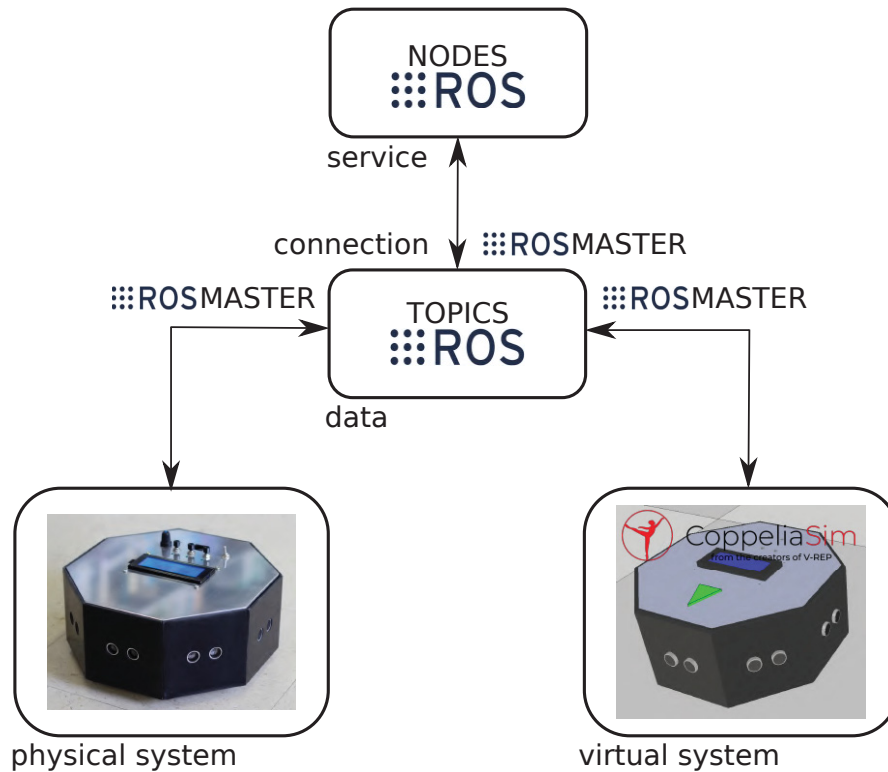


Figura 2.24: Digital Twin Mobile Robot concept model. The *physical* system, a real mobile robot in a real environment; *virtual system*, mobile robot model in a virtual scene; *service*, integration between the ROS nodes; *data*, exchange of ROS messages by topics; *connection*, data exchange in cyber-physical systems through the ROS master (Magrin et al., 2021).

2.7.1.1 Physical system (real world)

Using a real environment and building the mobile robot to apply the cognitive sensor fusion method helps better understand the real world. It indicates the best or worst features of each type of sensor for the perception-action cycle in an indoor and outdoor environment.

2.7.1.2 Virtual system (simulations)

The robot environment simulation helps create a virtual robot for testing locomotion, perception, and decision-making. Also, this makes creating *digital twins* possible with virtual robots and real robots, increasing the possibilities for testing algorithms, using different types of sensors, and interacting with other robots and environments for better results of the robotic system. The control system is trained primarily in simulation and then transferred to the real world. A robot simulator is an essential tool for robotic projects, and some simulators that integrate with ROS are presented, such as CoppeliaSim and Gazebo.

CoppeliaSim (<https://www.coppeliarobotics.com/>) is a robot simulator based on a distributed control architecture, with the integrated development environment used for fast algorithm development, factory automation simulations, and creating any robot. This framework can be controlled via a particular embedded script and a ROS node. One of the most important aspects is the simulation model's flexibility, portability, and scalability. CoppeliaSim

implements a ROS node via publishers and subscribers, and it's directly enabled from within the platform via an embedded script command. In virtual environments, including the most popular mobile robots, such as Khepera III, KUKA YouBot, NAO, and Pioneer 3-DX, designing the mobile robot model and simulation scene, which supports CAD data formats, is of great relevance for some projects of real robots and environments. The following scene objects are supported in CoppeliaSim: joints, shapes, proximity sensors, vision sensors, force sensors, cameras, lights, paths, and dummies.

Gazebo (<http://gazebo-sim.org/>) is an open-source simulator that makes it possible to test algorithms using ROS rapidly. This tool designs robots using SDF (Simulation Description Format), simulates robots in complex indoor and outdoor environments under various conditions, and provides many robots, including PR2, Pioneer 2-DX, iRobot Create, and TurtleBot. Gazebo generates sensor data from laser 2D/3D, Kinect, contact, and force sensors.

2.7.1.3 Service (ROS)

Integration between the ROS nodes enables them to communicate with other nodes. Services allow nodes to request and receive specific information on demand, supporting synchronous communication. This mechanism is useful for triggering actions that require confirmation or a particular response within the *cyber-physical* system.

2.7.1.4 Data (ROS)

Exchange ROS messages by topics, transmit messages among themselves through a common topic to *publish* data and *subscribe* data. Topics enable continuous and asynchronous data flow between nodes, supporting modular and scalable architectures. Such communication is fundamental for sharing mobile robotics sensor data, control commands, and system states.

2.7.1.5 Connection (ROS)

Data exchange and running programming use the same standard data and information in both *cyber-physical* systems through the ROS master. The ROS master manages the registration of nodes, enabling dynamic discovery and communication among distributed components. This infrastructure facilitates the integration of hardware, sensors, and algorithms in a unified robotic system.

2.8 CHAPTER CONSIDERATIONS

The main sensors used in mobile robotics, proprioceptive and exteroceptive, have been presented as a basis for studying sensor fusion. Siegwart et al. (2011) first introduces the perception-action cycle in autonomous mobile robots, and the scheme remains current. Still, the work's foundations address work in psychology, directing attention to the system at the knowledge level and the total cognitive system, encompassing perception, cognition, and the motor system. Cognitive systems inspired by human cognition and cognitive dynamic systems form the basis of studies relating the cyclic operation to perception, memory, attention, and intelligence. To conclude the study of the chapter, the perception-action cycle of the cognitive system is presented, based on the work of Haykin (2012) and Fuster (2005), with an analogy to the basic biological principle of an organism's dynamic adaptation to its environment. A brief introduction to the ROS framework and the concept of a digital twin mobile robot is presented, as these concepts are utilized in this work. The analogy with cognitive systems provides a conceptual foundation

for this work, guiding the design of a sensor fusion model inspired by biological principles to support localization and mapping in autonomous mobile robots.

3 RELATED WORK

Recent advances in autonomous mobile robotics, sensor fusion, and Simultaneous Localization and Mapping (SLAM) increasingly integrate cognitive methods, focusing on selective sensor fusion to weight heterogeneous sensor contributions for improved robustness and interpretability dynamically. These cognitive approaches extend to selective perception and brain-inspired decision-making, aiming to create more precise cognitive maps and enabling robots to localize in a hierarchical manner akin to human cognition. The evaluation of methodologies is dependent on comprehensive datasets. While significant contributions exist, including large-scale benchmarks for autonomous vehicles like KITTI and Oxford RobotCar, and hybrid environment datasets such as NCLT, a persistent and considerable deficiency remains in comprehensive datasets for wheeled mobile robots equipped with diverse low-cost multi-sensors operating in challenging hybrid indoor and outdoor environments. This gap is particularly evident concerning datasets that provide detailed geodetic ground truth for seamless transitions across these complex spaces.

3.1 RESEARCH DIRECTIONS

This section presents a detailed overview of recent studies and the approaches that have advanced the cognitive integration of sensors.

Selective sensor framework was proposed by Chandrasekaran and Conrad (2016), using Dempster-Shafer's theory of evidence to combine the grouped sensor data. In the system, all sensors are treated equally to achieve the best interaction, and only a few sensors are grouped depending on the context. The decision framework combines the data from multiple agents and improves the interaction process between the user and the robot. The interaction uses a sensor map to allocate slots to the sensors, and task maps, which hold task information, to develop a structure similar to the brain map to achieve human-like thinking when interacting with an actual human operator and performing the interaction.

Hierarchical efficient localization was proposed by Sarlin et al. (2018), using deep learning, a hierarchical approach similar to how humans naturally localize in a previously visited environment by first looking at the global scene appearance and subsequently inferring an accurate location from a set of likely places using local visual clues. In a hierarchical localization system, look for approximate locations at the map level and then attempt to estimate the precise pose.

Brain-inspired SLAM system, called RatSLAM, was proposed by Zhang et al. (2019) and is used to construct a cognitive map (the core of RatSLAM model) for a mobile robot, observing the information of odometry and visual scenes. The proposed approach employs a multi-sensor fusion method that can construct more precise cognitive maps for indoor environments. The method is different from most probabilistic SLAM systems. Brain-inspired SLAM uses pose cells, local view cells, and cognitive maps to simulate the rodent's spatial cognition mechanism. The multi-sensor fusion improves the accuracy of robot state estimation and represents the correct spatial relationship for constructing a cognitive map.

Selective sensor fusion was proposed by Chen et al. (2019), which fuses monocular images and inertial measurements to estimate the trajectory while improving robustness to real-life issues, such as missing and corrupted data or imperfect sensor synchronization. The selection process is conditioned on the reliability of the measurement and the dynamic environment. The feature weighting strategies implemented two alternatives (deterministic and stochastic) to

learn the most relevant feature representations. The results show that features extracted from different modalities (vision and inertial motion) are complementary in various conditions: the inertial features contribute more in fast rotation, while visual features are preferred during large translations.

Cognitive understanding of environments was proposed by Lee et al. (2020), and this approach enables mobile robots to classify and differentiate the unknown environment through image object recognition using the proposed Cognitive SLAM. TurtleBot3 mobile robot platform explores the mapping of the house layout (kitchen, living room, bedroom, study room, and toilet) using a Gmapping package running in ROS. To achieve multiple object recognition using the YOLOv3 (Real-Time Object Detection) library and classify the environment through the surrounding objects based on the output results of the object recognition, train an Artificial Neural Network (ANN).

Cognitive decision making, proposed by Naghshvarianjahromi et al. (2020) is based on a cognitive dynamic system (CDS), inspired by the human brain decision-making (*brain-inspired*), and utilized for situation understanding in non-Gaussian and nonlinear environments. The proposed algorithm is used to control the quality of service in long-haul fiber-optic links. The technique adaptively uses different focus levels, and virtual actions increase 43% data rate enhancement compared to 13% for simple CDS. The executive block uses reinforcement learning to apply cognitive actions and receive rewards from the dynamic environment.

Selective perception was proposed by Ramicic and Bonarini (2020), a computational approach that filters the cognitive load selectively, thus emulating the attention mechanism characteristic of human perception. The model uses genetic algorithms to evolve the most efficient attention filter mechanism that would give the agent an optimal perception of a specific environment by discriminating which experiences are valuable for learning. The proposed approach could develop a filtering mechanism to reduce the cognitive load and induce an effective intrinsic behavior.

Cognitive functions in robotics are recent works, delimited experiments regarding Piaget's theory about the construction of knowledge and intelligence development. The methodology used scenes to validate the Piaget stages in the cognitive attentional model. The cognitive architectures using reinforcement learning algorithms, usually Markov decision, evaluate the existence of cognitive mechanisms of perception, attention, reasoning, memory, and learning. Attention has the main purpose of selecting relevant information and filtering irrelevant information from sensory data (Berto, 2020; Rossi, 2021). The cognition architecture uses the platform Cognitive Systems Toolkit (CST) for the construction and validation of the cognitive functions (Paraense et al., 2016).

Sensor Fusion between indoor and outdoor environments was proposed by Tian and Mirza (2022) to develop an Octagon WMR and simulation model in ROS. The sensor fusion used for indoor navigation of autonomous robots, 2D LiDAR, and Kinect improves position estimation accuracy. Odometry with encoders and GPS fusion increases outdoor robot navigation. This article indicates the algorithm for changing indoor and outdoor environments during motion for future work development.

Cognition-aided reinforcement learning was proposed by Rathore and Bhadauria (2022), where the algorithm combines reinforcement learning with multiple cognitive principles for making intelligent driving decisions in autonomous vehicles. The study predicts future states based on multiple cognitive inputs, such as perception (multiple sensors), language (essential safety messages from other agents), memory, attention, and intelligence. Simulations show that the proposed algorithm is superior to the state-of-the-art. The research suggests that language

and the perception of prior experiences stored in the memory improve the quality of intelligent decision-making.

Learning Selective Fusion was proposed by Chen et al. (2025), a generic end-to-end framework for selective sensor fusion that learns to dynamically weight contributions to features of heterogeneous modalities, fusing a pair of modalities from vision, depth, inertial, and LIDAR data. Unlike traditional approaches that fuse all sensor features indiscriminately, SelectFusion incorporates deterministic (soft) and stochastic (hard) fusion modules to handle degraded or misaligned data. The method demonstrates superior performance across various state estimation tasks: visual-inertial odometry (VIO), LIDAR-vision odometry, and RGB-D relocalization on both public datasets and synthetically corrupted datasets. The ability to interpret learned feature masks offers valuable information on modality reliability under different motion and environmental conditions, thus guiding future sensor suite design and adaptive fusion strategies for mobile robots.

The research directions explored in this section highlight the growing convergence between cognitive science and robotic perception, emphasizing how selective, hierarchical, and learning-based fusion frameworks are shaping the next generation of autonomous systems. From brain-inspired mapping and cognitive decision-making to selective fusion strategies that dynamically adapt to environmental uncertainty, these advances demonstrate a progressive shift toward self-adaptive, perception-driven robotics. Collectively, these approaches reinforce the foundation for developing cognitive sensor architectures capable of reasoning, learning, and adapting.

3.2 DATASETS

This section presents the main datasets used to benchmark the research in wheeled autonomous mobile robots and autonomous vehicles. A summary of the papers is presented on the datasets with different data collection features or even the application of sensors. At the end of this section, Table 3.1 shows the most common datasets applying multi-sensors, comparing the types of sensors to analyze the sensors used in this work. Representations of sensor data for an autonomous mobile robot in outdoor environments have been released in autonomous vehicles research, including dataset development of Pandey et al. (2011); Blanco-Claraco et al. (2014); Sun et al. (2020), and with a focus on *The KITTI dataset* (Geiger et al., 2013), *The Cityscapes dataset* (Cordts et al., 2016), *The Oxford RobotCar dataset* (Maddern et al., 2017), and *PandaSet* (Xiao et al., 2021). The indoor environments, including the dataset development of *RGB-D SLAM dataset* (Sturm et al., 2012), and *North Campus Long-Term (NCLT) dataset* indoor and outdoor environment (Carlevaris-Bianco et al., 2016).

Despite advancements in mobile robotics research, the literature still presents a significant deficiency in comprehensive datasets for wheeled mobile robots equipped with multi-sensors in hybrid environments. In recent years, there has been an increased representation of datasets collected in indoor and outdoor settings, often utilizing robots found in research robotics laboratories. Notably, the *M2DGR dataset* (Yin et al., 2022) stands out for employing multiple camera types for data collection, while *TIERS dataset* (Qingqing et al., 2022) offers a rich source of LiDAR-based sensor data. Our UFPR-MAP dataset advances this scenario by incorporating a variety of low-cost sensors, which is crucial for the replication and accessibility of research. Furthermore, the primary differential of the UFPR-MAP lies in its capability to provide detailed geodetic ground-truth for hybrid indoor and outdoor environments. This feature represents a substantial advancement in precisely validating localization and mapping algorithms.

The *KITTI dataset* is a standard benchmark for odometry estimation, consisting of a collection of sensor data from autonomous vehicle scenarios (Chen et al., 2020). In addition, other vision tasks use the benchmark, such as stereo, optical flow, visual odometry, 3D object detection, and 3D tracking. The autonomous platform (Figure 3.1) is a VW Passat B6 modified with pedals and steering wheel actuators. The following sensors were used to record the dataset: inertial navigation system (GPS/IMU) Oxts RT 3003, laser scanner Velodyne HDL-64E, grayscale camera Point Grey Flea 2 with varifocal imaging lenses, and color cameras Point Grey Flea 2 with varifocal imaging lenses. The KITTI datasets were captured by driving around the mid-size city of Karlsruhe (Germany), in rural areas, and on highways (Geiger et al., 2013). According to Chen et al. (2020), the KITTI datasets are relatively simple and small. It is not convincing if only results on the KITTI benchmark are provided without a comprehensive evaluation in the real-world experiment. In other words, the dataset needs to create a benchmark for a thorough system evaluation covering various environments, self-motions, and dynamics.

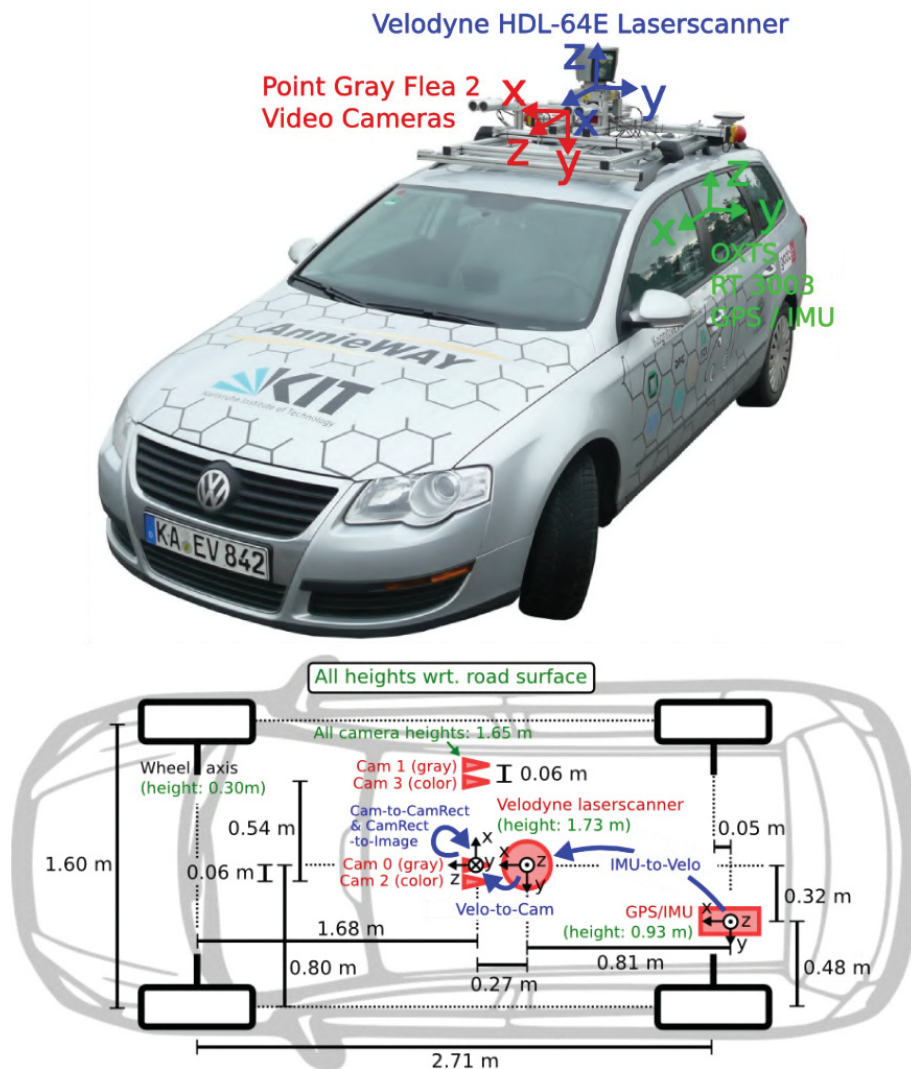


Figura 3.1: Sensors suite on the autonomous vehicle (VW Passat). The figure shows the multi-sensors in the vehicle, such as the inertial navigation system (GPS/IMU), laser scanner (Velodyne), and cameras (Point Grey Flea 2) (Geiger et al., 2013).

The *Cityscapes Dataset*, a large-scale dataset that contains a diverse set of stereo video sequences recorded in street scenes from 50 different cities, primarily in Germany but also in neighboring countries, in several months and adverse weather conditions (heavy rain or snow).

The dataset adequately captures the complexity of real-world urban scenes, intended to assess the performance of vision algorithms for major tasks of semantic urban scene understanding and to support research for training deep neural networks. The sensors were mounted behind the windshield and provided depth information through the stereo camera. The dataset also includes vehicle odometry, outside temperature, and GPS tracks (Cordts et al., 2016).

The Oxford RobotCar dataset, a large-scale dataset for autonomous vehicles and mobile robots approaches to localization and mapping, contains over 1000km of recorded driving in a consistent route through Oxford (UK), captured over a year. Sensors collected data in many different combinations of weather (heavy rain, night, direct sunlight, and snow), traffic, and pedestrians, along with longer-term changes such as construction and roadworks. The Oxford RobotCar platform, an autonomous Nissan LEAF (Figure 3.2), is equipped with a trinocular stereo camera Bumblebee XB3, monocular camera Grasshopper 2, 2D LiDAR SICK LMS-151, 3D LiDAR SICK LD-MRS, and inertial/GPS navigation system NovAtel SPAN-CPT ALIGN. The Oxford RobotCar Dataset documentation provides simple MATLAB development tools and Python functions for easy access to and manipulation of the dataset (Maddern et al., 2017).

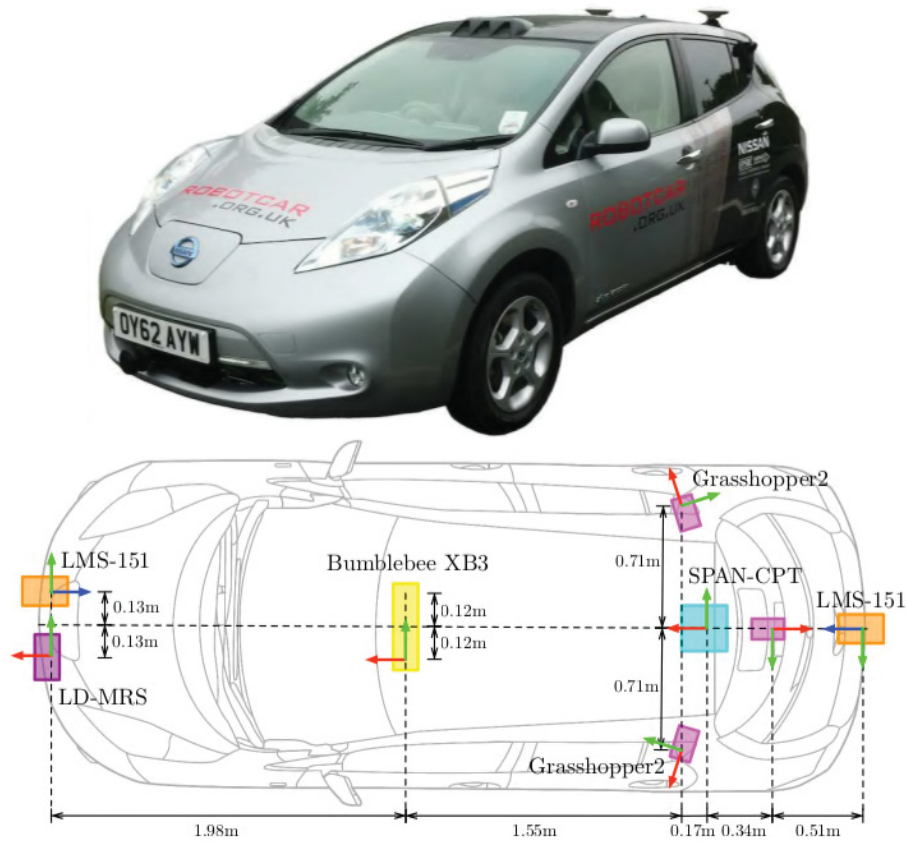


Figura 3.2: Sensors suite on the Oxford RobotCar. The figure shows the multi-sensors in the vehicle, such as trinocular stereo camera (Bumblebee), monocular camera (Grasshopper 2), 2D LiDAR (LMS-151), 3D LiDAR (LD-MRS), and inertial/GPS navigation system (SPAN-CPT) (Maddern et al., 2017).

The datasets *KITTI* and *Oxford RobotCar* were developed with the best sensors available on the market. According to Chen et al. (2020), a survey on “Deep Learning for Localization and Mapping”, the KITTI dataset is a common choice to evaluate visual odometry. Still, previous works split training and testing data differently, bringing difficulties to directly comparing them.

RGB-D SLAM Dataset is a large dataset containing RGB-D data and ground-truth data for evaluating visual odometry and SLAM systems. The dataset includes the color images, depth images, and the accelerometer data of a Microsoft Kinect sensor along the ground-truth trajectory

of the sensor. The system obtained the ground-truth trajectory from a motion-capture system with eight high-speed tracking cameras. The RGB-D SLAM was recorded from a Kinect mounted on top of a Pioneer robot, and the ROS bag file included the laser scan and the odometry data of the robot (Sturm et al., 2012).

The University of Michigan NCLT Dataset, a long-term autonomy dataset, was collected to facilitate robotics research focusing on long-term autonomous operation in changing environments and provide ground-truth pose in a single frame of reference. The Segway robot used to collect data is outfitted with a Ladybug3 omnidirectional camera, a Velodyne HDL-32E 3D LiDAR, Hokuyo planar LiDAR, an inertial measurement unit (IMU), and a global positioning system (GPS). The sessions explore the indoor and outdoor environments of the University of Michigan campus on varying trajectories and at different times of the day across all four seasons. The dataset captures many elements, including buildings, infrastructures, trees, cars, people, and others (Figure 3.3). NCLT provides the sensor data in CSV format and support for generating ROSbags (Carlevaris-Bianco et al., 2016).

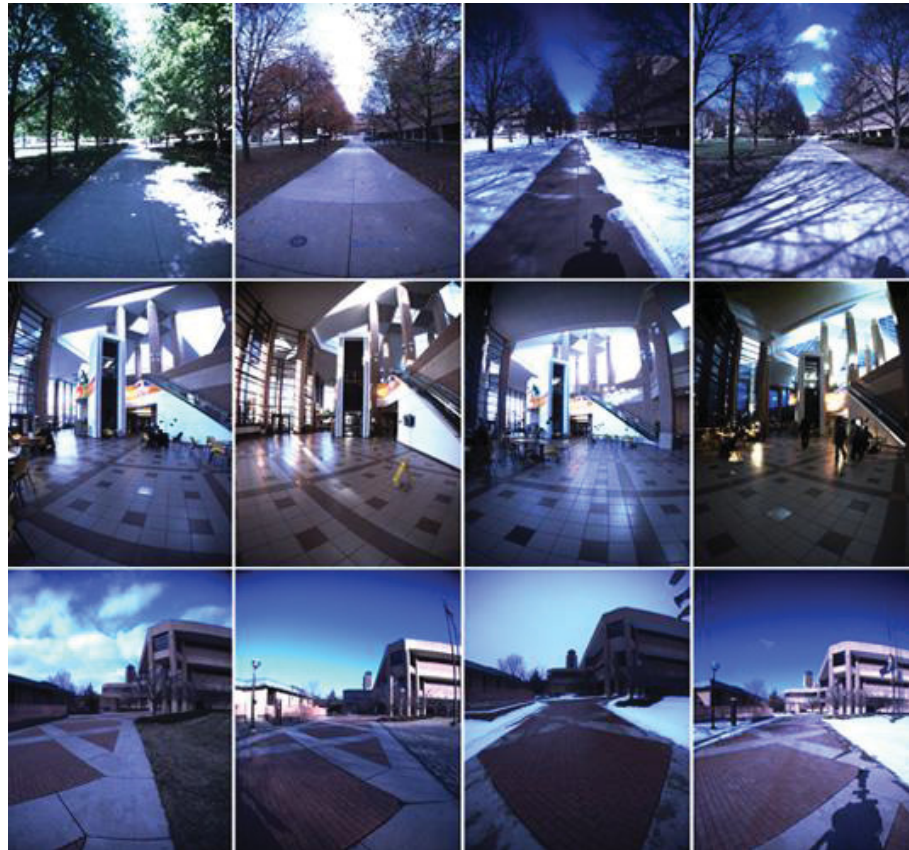


Figura 3.3: Sample imagery from the NCLT dataset. The figure shows samples in indoor and outdoor environments and at different times of the day (Carlevaris-Bianco et al., 2016).

According to Sarlin et al. (2018), the NCLT is one of the rare datasets, containing a large set of images captured at the exact geographical locations in perceptually changing conditions and accurate ground truth pose for every image. The NCLT is one of the few datasets that explores information from indoor and outdoor environments.

Table 3.1 shows the most common datasets that use multi-sensors, are organized chronologically, and present the platform type, the multi-sensors used, and the ground truth. Considering the most used datasets, they all use larger platforms, mostly cars, and with the best sensors found on the market, such as Teledyne FLIR and Hokuyo Automatic.

Tabela 3.1: The most common multi-sensors datasets. The table was organized in dataset chronological order: Ford Campus (Pandey et al., 2011), KITTI (Geiger et al., 2013), NCLT (Carlevaris-Bianco et al., 2016), Málaga Urban (Blanco-Claraco et al., 2014), and Oxford RobotCar (Maddern et al., 2017).

Dataset	Platform	Multi-sensors	Ground-truth
Ford Campus (2009)	Ford F-250 car	1x Velodyne HDL-64E LiDAR, 2x Point Grey Ladybug3 camera, 2x RIEGL LMS-Q120 LiDAR, 1x Applanix POS-LV 420 INS with GPS, 1x Xsens MTi-G MEM IMU	Navigation system (GPS and IMU) Environment: Outdoor
KITTI (2012)	VW Passat car	4x Point Grey Flea2 cameras, 1x Velodyne HDL-64E LiDAR, 1x OXTS RT3003 IMU and GPS.	Navigation system (RTK GPS and IMU) Environment: Outdoor
NCLT (2013)	Segway	1x Velodyne HDL-32E LiDAR, 1x Point Grey Ladybug3 camera, 1x Hokuyo UTM-30LX LiDAR, 1x Hokuyo URG-04LX LiDAR, 1x Microstrain 3DM-GX3-45 IMU, 1x KVH DSP-3000 single-axis FOG YAW, 1x Garmin GPS 18x 5Hz, 1x NovAtel DL-4 plus RTK GPS	Navigation system (RTK GPS and IMU) Environment: Outdoor/Indoor
Málaga Urban (2014)	Citroen C4 car	1x Point Grey Bumblebee2 camera, 1x Xsens MTi-G MEM IMU 3x Hokuyo UTM-30LX LiDAR, 2x SICK LMS-200 LiDAR, 1x GPS (low-cost).	Low-cost GPS Environment: Outdoor
Oxford RobotCar (2015)	Nissan LEAF car	1x Point Grey BumblebeeXB3 cam., 3x Point Grey Grasshopper2 cam., 2x SICK LMS-151 2D LiDAR, 1x SICK LD-MRS 3D LiDAR, 1x NovAtel SPAN-CPT IMU and GPS.	Navigation system (GPS and IMU) Environment: Outdoor

Research was conducted over the last 5 years to assess the trend in mapping sensors used in ground mobile robotics as benchmarks. Table 3.2 shows a trend in recent years toward mobile robotics platforms for use in indoor and outdoor environments and using various types of sensors. Datasets such as M2DGR (Figure 3.4) are aimed at cameras, and TIERS (Figure 3.5) are aimed at LiDAR. Concern with ground-truth to meet the transition between hybrid environments also follows as a line of research, depending on the environment and the length of the path traveled for mapping.

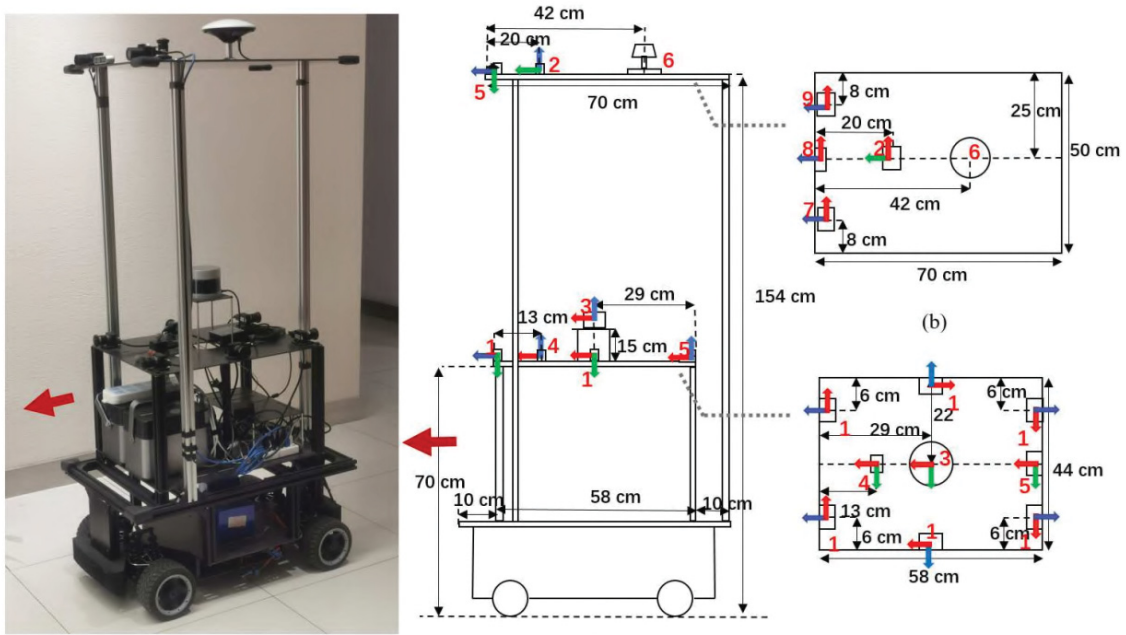


Figura 3.4: Sensors suite on the M2DGR robot. The figure shows the multi-sensors in the wheeled mobile robot, such as 2 - monocular camera (FLIR Chameleon3), 3 - 3D laser scanner (Velodyne VLP-32C), 4 - GNSS receiver (Ublox M8T), 5 - IMU (Handsfree A9), 7 - infrared camera (Gaode PLUG 617), 8 - depth camera (Intel Realsense d435i), and 9 - event camera (Inivation DVXplorer) (Yin et al., 2022).

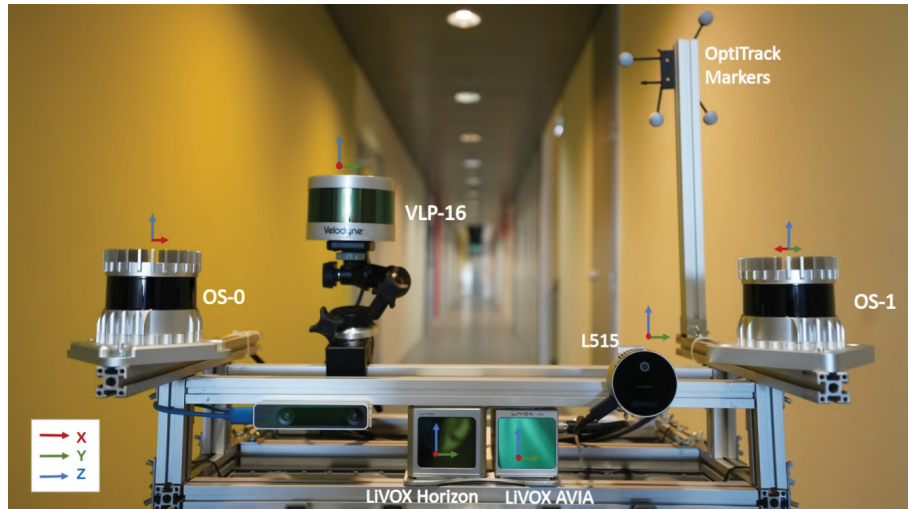


Figura 3.5: Sensors suite on the TIERS robot. The figure shows the multi-sensors in the wheeled mobile robot, such as a 3D laser scanner (Velodyne VLP-16), 3D laser scanner (Ouster OS1-64), 3D laser scanner (Ouster OS0-128), 3D laser scanner (Livox Horizon), 3D laser scanner (Livox Avia), and LiDAR depth camera (Intel RealSense L515) (Qingqing et al., 2022).

3.3 CHAPTER CONSIDERATIONS

Some works directed the proposed method, such as the combination of a set of sensors, construction of cognitive maps, or even the application of decision-making to control the quality of service of a fiber optic link by applying cognitive actions using reinforcement learning, as well as works indicating the exchange of environments in indoor and outdoor navigation and filter for cognitive load selection in the selective perception process. The research on the main datasets resulted in a comparative table of the most common types of sensors, such as LiDAR, camera,

Tabela 3.2: Main multi-sensor datasets in the last 5 years, highlighting the trend of using wheeled mobile robots and hybrid environments. The table is organized in dataset chronological order: PandaSet (Xiao et al., 2021), M2DGR (Yin et al., 2022), and TIERS (Qingqing et al., 2022).

Dataset	Platform	Multi-sensors	Ground-truth
PandaSet (2021)	Chrysler Pacific car	6x Leopard AR023ZWDRB camera, 1x Hesai Pandar64 LiDAR, 1x Hesai PandarGT LiDAR, 1x NovAtel PwrPak7 GNSS. and IMU.	Navigation system (GPS and IMU) Environment: Outdoor
M2DGR (2022)	M2DGR wheeled robot	1x Velodyne VLP-32C LiDAR, 1x Handsfree A9 IMU, 1x Ublox M8T GNSS, 1x Point Grey Chameleon3 camera 1x Gaode PLUG 617 infrared cam., 1x Inivation DVXplorer event cam., 1x Intel Realsense d435i depth cam.	Navigation system (RTK/INS) Total station Environment: Indoor/Outdoor
TIERS (2022)	TIERS wheeled robot	1x Velodyne VLP-16 LiDAR, 1x Ouster OS1-64 LiDAR, 1x Ouster OS0-128 LiDAR, 1x Livox Horizon LiDAR, 1x Livox Avia LiDAR, 1x Intel RealSense L515 LiDAR.	6DOF MoCAP and SLAM Environment: Indoor/Outdoor

IMU, and GPS. The datasets found in the literature typically use cars for outdoor mapping, and only the NCLT has an adaptation of a Segway for indoor mapping in addition to outdoors. Recent research has shown a different trend for datasets utilizing wheeled mobile robots and employing an approach that uses camera and LiDAR sets for sensor mapping. There is a gap in building a georeferenced dataset with an indoor-to-outdoor transition using a mobile robot with multi-sensors.

4 THE PERCEPTION-ACTION CYCLE

This chapter presents the perception-action cycle of an autonomous mobile robot cognitive sensor fusion process, presenting an analogy with a human's basic cognitive perception-action cycle shown in Figure 2.21. The proposed sensor blend sets reinforce the cognitive system process as an alternative to improve localization and navigation in hybrid environments. The cognitive blended sensors method is detailed using the sensor blend sets in the sensor fusion process, directing cognitive learning.

4.1 COGNITIVE SENSOR FUSION

The goal of sensor fusion is the perception of many environment features, explored by proprioceptive and exteroceptive sensors. *Proprioceptive sensors* measure robot parameters with typical features, such as acceleration, motor speed, position, motor load, and robot orientation about a fixed reference. *Exteroceptive sensors*, extract features from the robot in the environment, with general typical use, such as closeness, distance, light intensity, the orientation of the robot about a fixed frame, localization in a fixed frame, geometric triangulation, visual ranging, semantic image segmentation, and object recognition.

We propose identifying the cyclic operation of the cognitive system for mobile robot sensor fusion in self-localization and exploring cognitive processes, such as the perception-action cycle, memory, and attention.

4.1.1 Perception-action cycle

The perception-action cycle in a wheeled mobile robot self-localization system, shown in Figure 4.1, observes the environment through the sensory hierarchy and controls the mobile robot in the environment through the motor hierarchy.

Sensory hierarchy, observe the surrounding environment in three hierarchical levels inspired by the sensory cortex. In cognitive blend sensors, each perception stage represents a sensory level: proprioceptive sensors, exteroceptive sensors, and sensor fusion.

- *Proprioceptive sensors* is the first self-sufficient stage of perception, **measuring robot parameters**, using sensors, such as an accelerometer, encoder, and gyroscope.
- *Exteroceptive sensors* is the second self-sufficient stage of perception, **extracting features from the robot in the environment**, using sensors, such as bumpers (switches), digital compass, GPS, RF sensor, ultrasound, infrared, laser scanner, and vision.
- *Sensor fusion* is the last stage of perception, **the integration of information** through proprioceptive and exteroceptive sensors. This sensor fusion stage uses an Interacting Multiple Model (IMM), which uses different models to improve the system's precision.

Motor hierarchy, controls the mobile robot in the environment, inspired by the motor cortex. Action stages are represented by each level of the motor, such as locomotion, motor control, and path planning.

- *Locomotion* is the first self-sufficient stage of actions, **execution of movement**, using PWM (Pulse Width Modulation) control to move the robot in the environment.

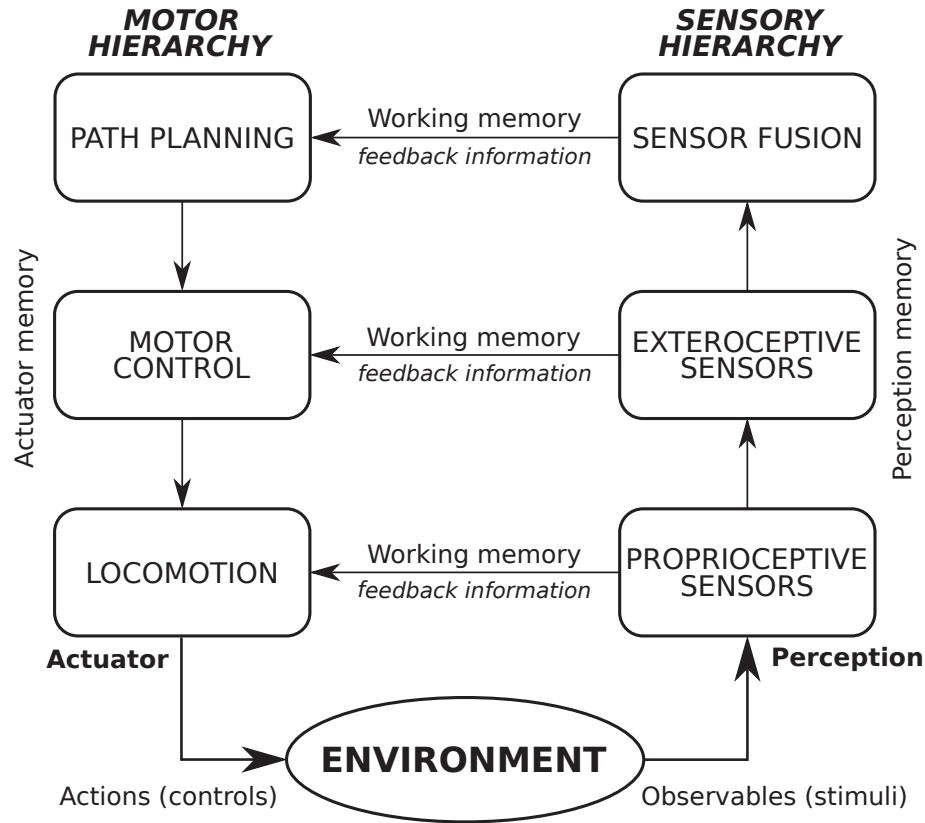


Figura 4.1: Perception–action cycle of a cognitive sensor fusion. The sensory hierarchy observes the surrounding environment in three hierarchical levels: proprioceptive sensors, exteroceptive sensors, and sensor fusion. The motor hierarchy controls the mobile robot in the environment, involving locomotion, motor control, and path planning. Working memory represents the state of the mobile robot; actuator memory means the action from the motor hierarchy, and perception memory means the action from the sensory hierarchy through a sensor fusion process. Inspired from (Haykin, 2012; Fuster, 2005).

- *Motor control* is the second self-sufficient stage of actions, **direct control of motor behavior**, using a PID controller (Proportional-Integral-Derivative) through a closed-loop with feedback information from proprioceptive and exteroceptive sensors.
- *Path planning* is the last stage of actions, **control executive functions**, decision making of path planning, the autonomous mobile robot, using artificial intelligence.

4.1.2 Memory

The data fusion of the sensory hierarchy learns from the environment using information through proprioceptive and exteroceptive sensors, and each sensor represents a feature of the environment. The sensor data continually change over time following changes in the environment, according to the received motor parameter feedback information in working memory, decisions about the mobile robot's locomotion in the environment change with time. Feedback information defines the error between the actual state of the environment and its estimate by the perception process. The actuator memory includes the actions from the motor hierarchy, and the perception memory consists of the actions from the sensory hierarchy.

4.1.3 Attention

Attention is a cognitive process for resource allocation and avoids information overload; in other words, the attention process filters information received from multi-sensors at the sensor fusion level. According to Fuster (2005) cognitive functions are sensory and motor interactions with the environment. Perceptual processing is guided by selective attention. In cognitive sensor fusion, attention is allocated selectively to the mobile robot localization.

4.2 SENSOR BLEND SETS

The sensor combination model, sensor blend sets (SBS), was proposed as an alternative to the state-of-the-art in sensor fusion (Magrin and Todt, 2024). The SBS meets the sensor fusion level of cognitive blend sensors. The robot can use the sensor blend sets to perform better in localization and navigation in indoor and outdoor environments.

Figure 4.2 shows the set of sensors respecting information from the cognitive process, starting with proprioceptive and simple contact sensors (bumper). Exteroceptive sensors provide complementary information for the sensor fusion process. The combination of sensor sets was segmented into five levels, divided into different features to complement the process of a cognitive system, and they are:

- **SBS-State** (battery current, gyroscope, and bumper): proprioceptive and contact sensors to check the robot's basic activities, such as battery charge, robot tilt, and collision (contact switch).
- **SBS-Movement** (accelerometer and encoder): the set with proprioceptive sensors has features for odometry and acceleration of the robot.
- **SBS-Distance** (ultrasound, infrared, and digital compass): set of distance measurement sensors and orientation.
- **SBS-Mapping** (laser scanner and vision): a set of sensors with information for mapping.
- **SBS-Positioning** (GPS and RF sensor): set of positioning sensors.

The choice of sets followed the similarity between the features and applications of the sensors in a cognitive sensor fusion. SBS-State, the set of sensors, indicates the battery status, collision with a specific object, and whether the robot is on the plane. SBS-Movement, the set of sensors, provides complementary information to estimate acceleration and distance traveled. SBS-Distance, the set of sensors, provides information to measure the distance in centimeters between the robot and objects with or without orientation invariance, in addition to providing the robot's orientation in the environment. Using multiple ultrasounds and multiple IRs provides a set of complementary and competitive information. SBS-Mapping, the set of sensors, provides complementary and cooperative information to build a map of the environment. Computer vision also allows you to identify the type of environment and objects in the scene. SBS-Positioning, the set of sensors, provides cooperative positioning information for indoor and outdoor environments; for example, GPS includes information for outdoor positioning and RF sensors for indoor environments.

The next section uses SBS to assist the proposed cognitive blended sensors method. Combining sensors with complementary, competitive, and cooperative information directs the level of sensor fusion and cognitive process tasks.

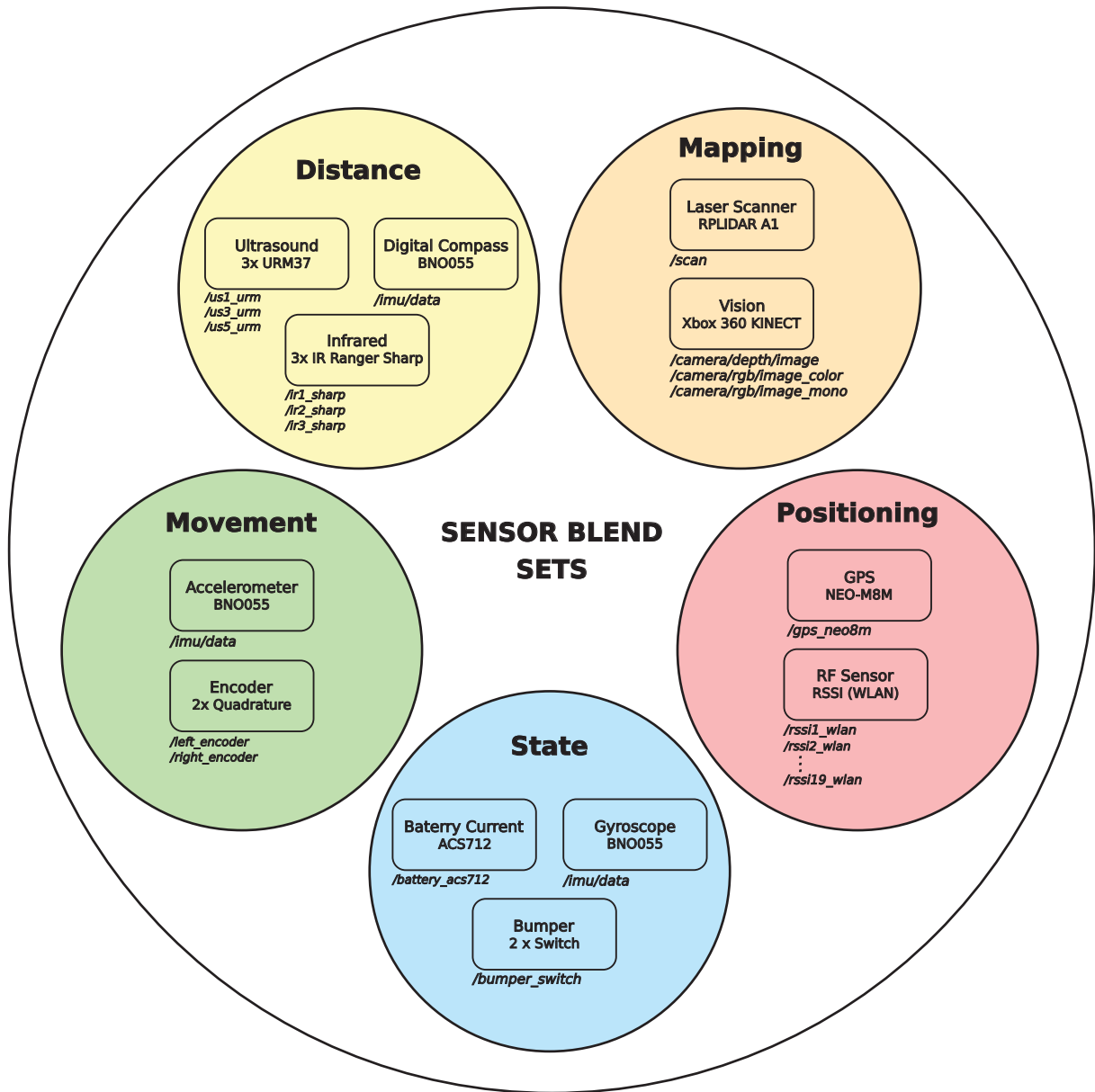


Figura 4.2: Sensor Blend Sets (SBS). Each of the five levels has a feature that combines and complements the process of a cognitive system. SBS-State (battery current, gyroscope, and bumper); SBS-Movement (accelerometer and encoder); SBS-Distance (ultrasound, infrared, and digital compass); SBS-Mapping (laser scanner and vision); and SBS-Positioning (GPS and RF sensor).

4.3 COGNITIVE BLENDED SENSORS METHOD

We proposed a Cognitive Blended Sensors (CBS) method to assist autonomous mobile robots in localization and mapping. Figure 4.3 shows the perception-action cycle of the proposed CBS in each sensory and motor hierarchical level. The SBS shown at level 2 is the same as that shown in Figure 4.2.

- **Environment:** dataset or robot moving in a real environment.
- **Level 0** (proprioceptiveSensors-locomotion): inspired by the sensory and motor cortex of the human brain, related in *primary sensory cortex* is dedicated to the representation of an elementary sensory feature, analog in mobile robot *proprioceptive sensors* check

battery charge (reference voltage), if the robot is on a plane (gyroscope), and robot speed (accelerometer and encoder). *Primary motor cortex* generates neural impulses that control the execution of movement, analogous in mobile robots to the execution of locomotion with PWM (Pulse-Width Modulation).

- **Level 1** (exteroceptiveSensors-motorControl): inspired by the sensory and motor cortex of the human brain, related in *unimodal association cortex* is elementary level in the hierarchy dedicated to the processing of complex information, analog in mobile robot *proprioceptive sensors* check for collisions through bumpers, orientation (digital compass), outdoors localization in latitude and longitude with GPS, indoor localization using RSSI attenuation through access points available in the environment (RF sensors), sonar octagon for reading 360-degree distance up to 4 meters, LiDAR for 360-degree environment mapping and vision exploring location recognition, semantic localization (Kinect and stereo cameras), recognizing environments such as indoor and outdoor, or even library, classroom, bathroom, or hallway. *Premotor cortex* direct control of motor behavior, playing a role in planning movement and using rules to perform specific tasks, analogous to the mobile robot *motor control* controller of the movement behavior using PID (Proportional-Integral-Derivative), according to the response of the proprioceptive and exteroceptive sensors.
- **Level 2** (sensorFusion-pathPlanning): inspired by the sensory and motor cortex of the human brain, *polymodal association cortex*, integration of information coming in through different senses, similar in *sensor fusion*. *Prefrontal cortex* controls executive functions, such as planning and decision making. The *path planning* position of the robot (localization) is input to motion planning. Sensor Blend Sets (Figure 4.2) - a combination of sensor sets divided into different features complements the sensor fusion process.
- **Attention** (cognitiveLearning): proposal for selecting the best combined sensor sets to assist in the mobile robot's perception, localization, and mapping process in hybrid indoor and outdoor environments.

4.4 CHAPTER CONSIDERATIONS

The proposed method initially identified the perception-action cycle of a cognitive system for autonomous mobile robots, suggesting a cognitive sensor fusion architecture of a perception-action cycle with a hierarchy of levels: proprioceptive sensors (measuring robot parameters); exteroceptive sensors (extracting characteristics from the robot in the environment); and sensor fusion (integrating information). The action cycle with a motor hierarchy is represented by three levels: locomotion (execution of movement), motor control (direct control of motor behavior), and path planning (control of executive functions). As an alternative to assist sensor fusion, sensor blend sets were proposed, directing the process of combining different types of sensors for each task in the cognitive system process: state, movement, distance, mapping, and positioning. The cognitive blended sensors (CBS) method results from the proposed perception-action cycle of the cognitive system for mobile robots and the sensor blend sets at the sensor fusion level. The CBS perception-action cycle learning process can occur in the last layer, encompassing attention and selecting the best sensor blend sets. Experiments are guided by these sensor blend sets, aligning the organization of topics within this work with the hierarchical structure of the

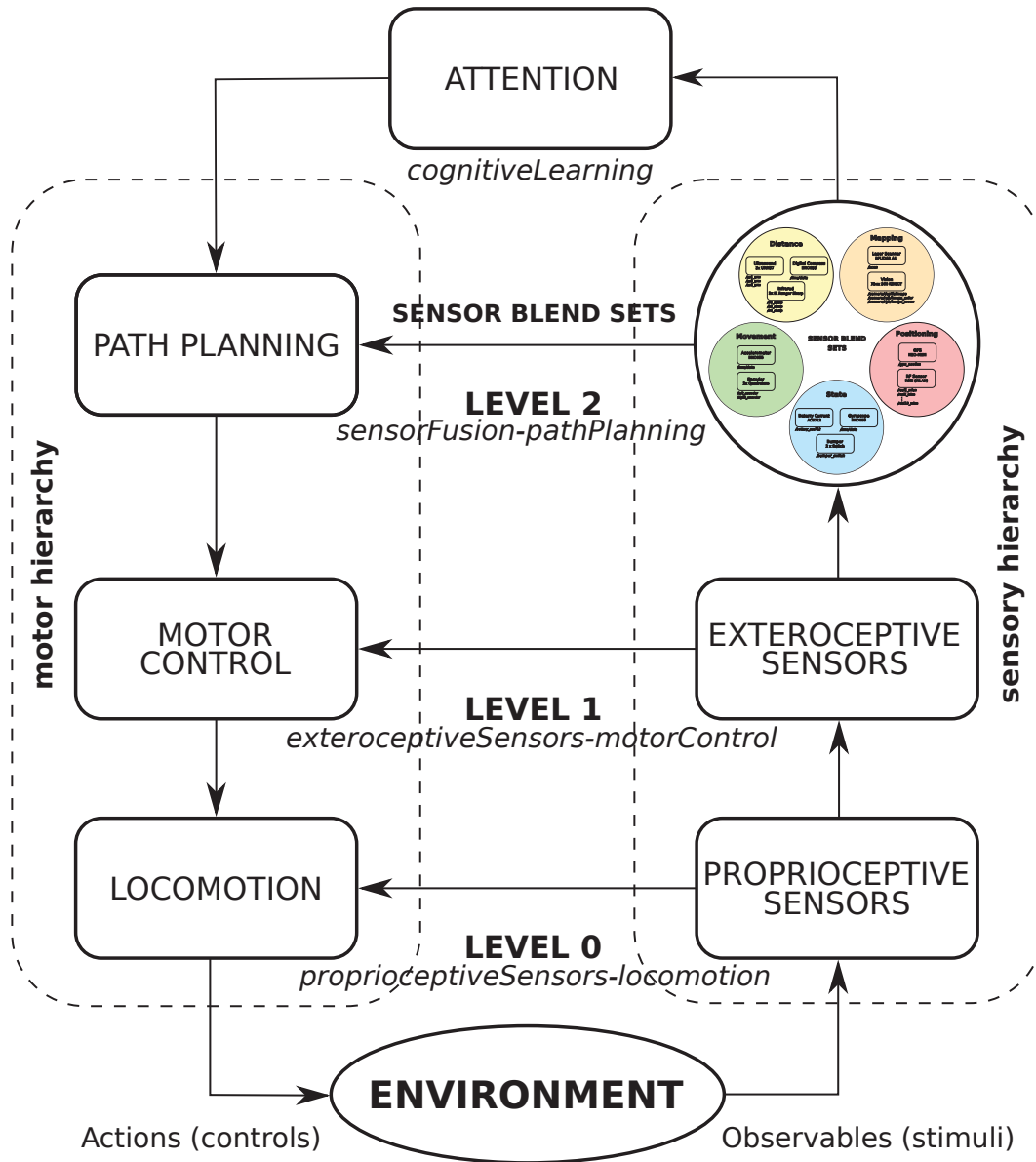


Figura 4.3: Learning perception-action cycle of cognitive blended sensors in each level. The figure shows the hierarchical levels of sensory and motor, such as environment, proprioceptive sensors and locomotion (level 0), exteroceptive sensors and motor control (level 1), sensor fusion and path planning (level 2 - SBS), and cognitive learning (attention) selection of the best sensor blend sets.

cognitive blend sensors. The research on cognitive systems and, consequently, the proposed perception-action cycle, motivated the development of the mobile robotics platform dedicated to mapping multi-sensors in indoor and outdoor environments, is presented on Chapter 5. The construction of the dataset, using raw sensor data and publishing in ROS messaging, contributes to the robotics community's learning process for sensor fusion in hybrid environments using a wheeled mobile robotics platform, is presented in Chapter 6.

5 MOBILE ROBOT PLATFORM

The chapter begins by describing the early experiments conducted in simulation environments and with existing datasets, and demonstrating initial validations of essential concepts for starting to build the mobile robot platform. The principles from these early experiments directly shaped the subsequent development of the VRI4WD platform, influencing its design, hardware specifications, and sensor suite, while emphasizing its suitability for hybrid environment locomotion. The development of the robotic platform was dedicated to mapping based on low-cost sensors, aiming to build an innovative dataset. This dataset is presented in Chapter 6. To address this task, taking the cognitive blended sensors model as a guide, the VRI4WD mobile robot platform was custom-built as a dedicated tool to study advanced sensor fusion concepts, specifically the proposed sensor blend sets (SBS) model.

5.1 EARLY EXPERIMENTS

The early experiments were performed by building a maze environment in a Gazebo simulator to test SLAM using the ROS Gmapping package, based on Rao-Blackwellized particle filter (RBPF) occupancy grid mapping. A *digital twin* was created in CoppeliaSim to flexibly develop and validate the proposed method with the real model of the robot and the scene of the Pinheirinho campus at the University of Curitiba environment. The virtual model was developed in the CoppeliaSim simulator, integrated with ROS, to validate the sensor's perception and robot locomotion in the environment. We validated the algorithms for mapping, classification, object recognition, and learning in Python and C.

5.1.1 Simultaneous Localization and Mapping

For building a map using SLAM, a basic robot with odometry data and a generic laser scanner mounted on the top of the robot was simulated in Gazebo, navigating through a maze environment. Figure 5.1 shows the maze environment of the autonomous mobile robots championship organized by the Robotics and Embedded Systems (ROSIE) research laboratory, used as a model for evaluating SLAM in Gazebo. According to Santos et al. (2013), the Gmapping SLAM algorithm is the most widely used SLAM package in robots worldwide. Our SLAM evaluation uses the Gmapping package as a ROS node, creating a 2-D occupancy grid map from the laser and pose data of the mobile robot. We ran SLAM on the differential drive robot, following the launch files procedure described in (Joseph, 2018).

Figure 5.2 shows the simultaneous localization and mapping in a *RViz*, a 3D visualization tool for ROS, simulating locomotion in a maze Gazebo scene with a one-meter grid and measuring 6m by 6m. The Gmapping package creates map files similar to the environment for localizing mobile robots and path planning on a static map. The map server package stored the YAML file, which contains the map metadata, and the image PGM file (Figure 5.3), which has the encoded data of the occupancy grid map.

5.1.2 Mobile Robot Model - TUPY 4WD

We developed a mobile robot for this work that applies multi-sensors and locomotion in indoor environments. The first design was a TUPY-4WD robot (Figure 5.4). We set the mobile robot to include and adjust many types of sensors. In CoppeliaSim, a *digital twin* of both the

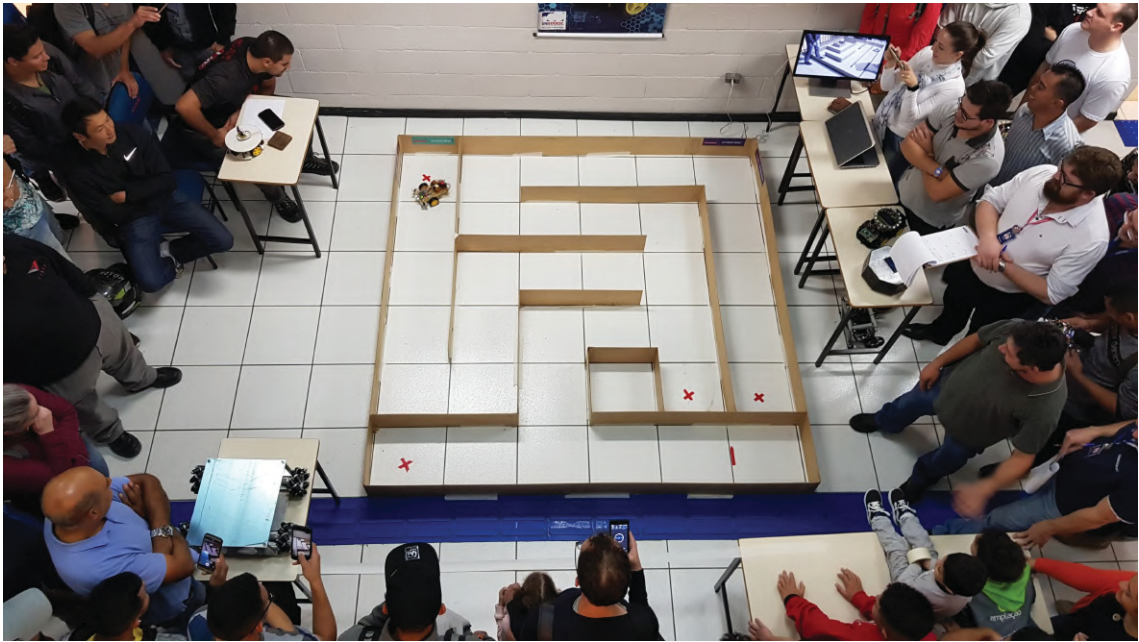


Figura 5.1: Maze environment of the Autonomous Mobile Robots Championship organized by the Laboratory ROSIE/UNICURITIBA. We used it in the design of the scene in Gazebo.

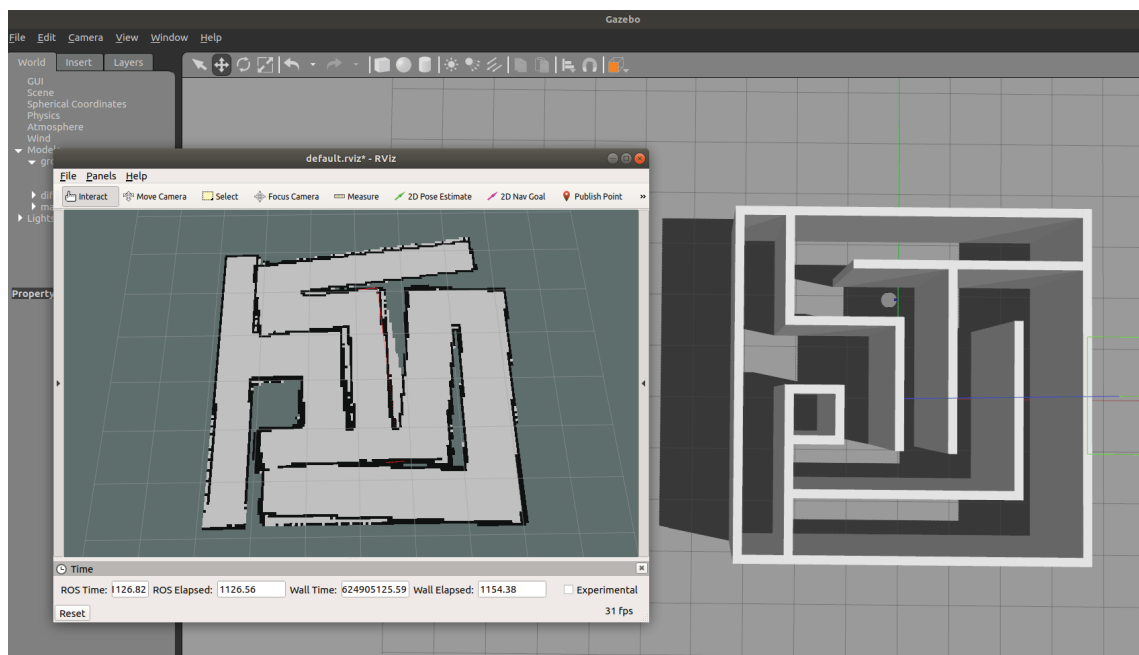


Figura 5.2: The Gmapping SLAM package shows in RViz and simulates locomotion in a maze Gazebo scene. Using a differential drive mobile robot in Gazebo, a 2-D occupancy grid map from the pose data and laser scanner mounted on top of the mobile robot was created.

environment and the robot was created, and ROS controlled the tasks using an embedded script. Figure 5.5 shows the TUPY-4WD robot model in the test with different types of sensors available in CoppeliaSim, such as proximity sensor (eight sonars in the octagon), Kinect (vision), Hokuyo (laser scan), reference frame (odometry), and RF sensor (Euclidean distance between the robot from the objects “access point” in the scene).

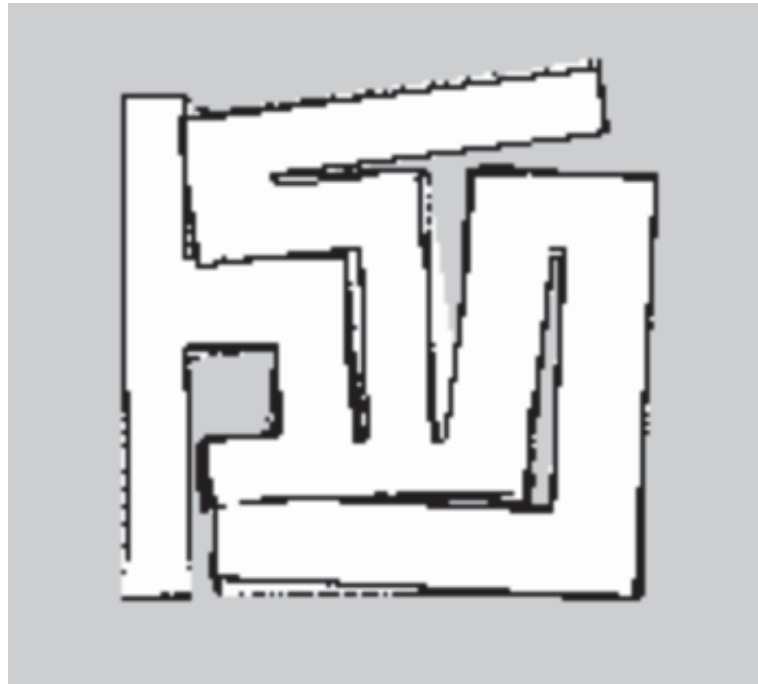


Figura 5.3: The image map, which has the encoded data of the occupancy grid map, is saved using the map server package. The map server package stored the YAML file containing the map metadata and the image PGM file.

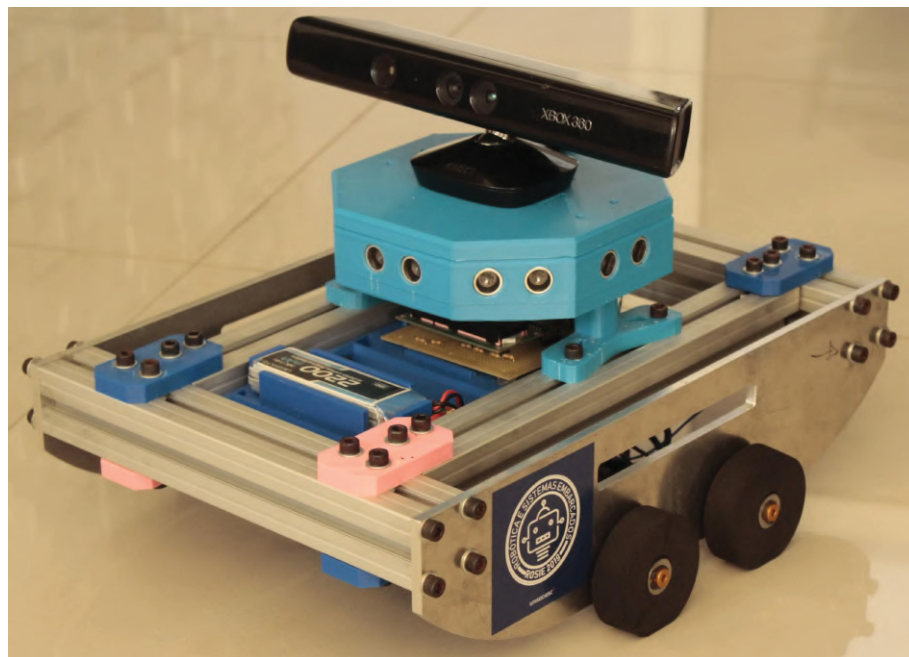


Figura 5.4: The physical TUPY-4WD robot was developed to apply multi-sensors and locomotion in indoor environments and is set to include and adjust many types of sensors.

5.1.3 Creating a Scene - Pinheirinho Environment

To create the virtual environment (scene), the Pinheirinho campus basement floor plan at the University of Curitiba (UNICURITIBA) was used as the real model. The mobile robot TUPY 4WD with the CAD project was developed for robotic learning at the University. Figure 5.6 shows a photo of the Pinheirinho campus.

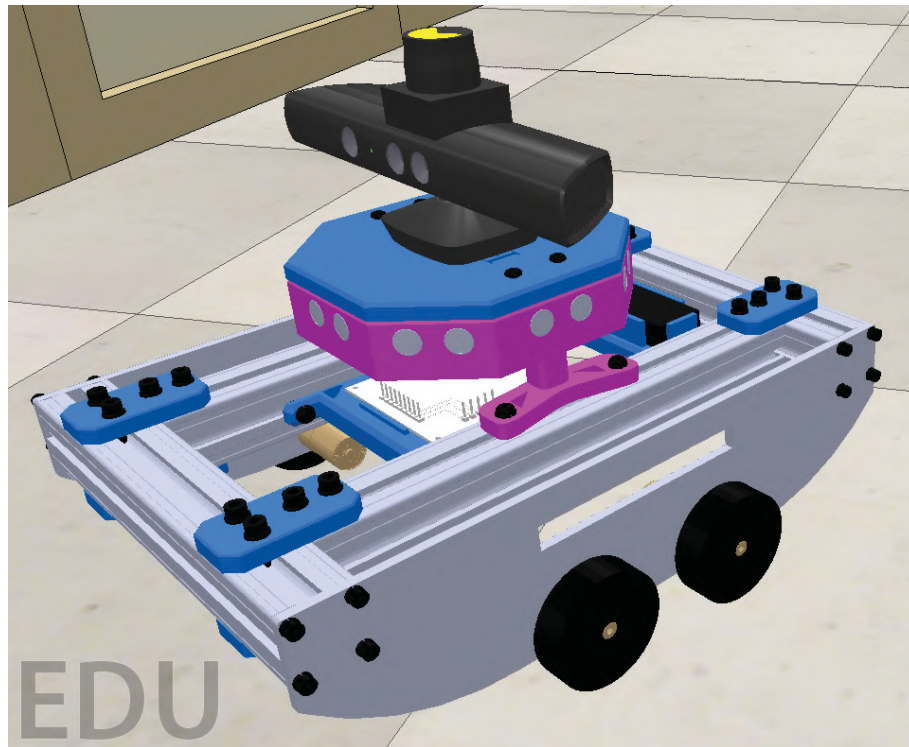


Figura 5.5: The virtual TUPY-4WD robot in Coppeliasim. The figure shows mounted on the top of the mobile robot eight ultrasound sensors in an octagon, a Kinect, and a Hokuyo laser scanner (Magrin et al., 2021).



Figura 5.6: Pinheirinho campus at the University of Curitiba. We used the floor plan to model indoor and outdoor environments in Coppeliasim. The figure shows the vehicle parking and entrances 'C' block and 'D' block.

The Pinheirinho virtual environment was created using infrastructure available in the Coppeliasim (walls, floors, doors, and others), and robot model export STL (Standard Triangle Language) file format using *SolidWorks to URDF Exporter* (http://wiki.ros.org/sw_urdf_exporter), representing a TUPY 4WD Robot (base, sensors, and joints). The exporter can automatically determine the proper joint type, joint transforms, and axes. The Unified Robot Description Format (URDF) is an XML file format used in ROS to describe all elements of a

robot. Figure 5.7 shows the virtual scene and the TUPY 4WD robot model simulation locomotion and perception of the indoor environment on CoppeliaSim.

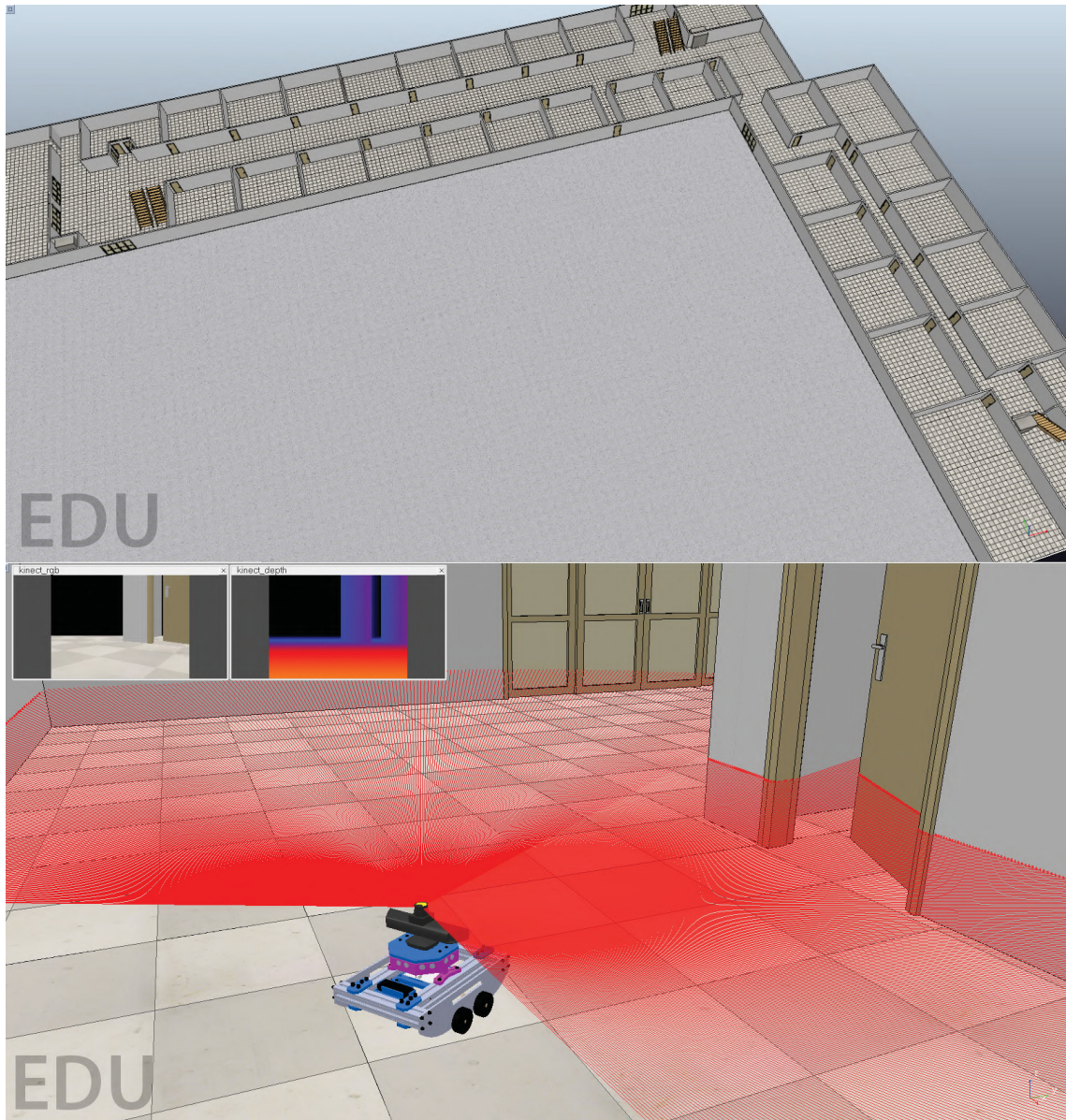


Figura 5.7: CoppeliaSim indoor and outdoor virtual environment of the Pinheirinho Campus at the University of Curitiba. Simulation TUPY-4WD robot model for locomotion and perception of the indoor environment, using a laser scanner and depth image/RGB image from Kinect.

5.1.4 Locomotion and Perception - ROS interface

In this experiment, the interface simulation using ROS provides connection information to *nodes*, and transmits messages among them, through a common topic to *publish* and *subscribe* data. ROS visualizes the computational graph using the package *rqt_graph*. Figure 5.8 shows information *nodes*, such as */csf_robotCS_teleop* - robot teleoperation *publish* topic */cmd_vel*, */sim_ros_interface* - general ROS functionality in CoppeliaSim, */csf_robotCS_cannyKinect* - package *CvBridge* converts between ROS Image messages and OpenCV images, Canny is a popular edge detection algorithm and OpenCV function *subscribe* topic */kinect/rgb_image* for edge detector output, and */csf_robotCS_perception* - preprocessing sensors information, topics

subscribe */poseAP*, */sonar*, */positionRobot*, and *orientationRobot* for publish sonars with data invariant orientation, mobile robot orientation by yaw axis, and Euclidean distance between the robot from the objects “access point” in the scene.



Figura 5.8: Displaying graphs of running ROS nodes (ellipses) with connecting topics (rectangles) and packages (obtained by the tool *rqt_graph*). The ROS nodes running robot teleoperation */csf_robotCS_teleop*, interface ROS with CoppeliaSim */sim_ros_interface*, edge detection algorithm and OpenCV function with Kinect */csf_robotCS_cannyKinect*, and preprocessing sensor information */csf_robotCS_perception*.

The mobile robot performed the testing in the CoppeliaSim simulator with ROS interface. Figure 5.9 shows Hokuyo (proximity sensor - laser) scanning the scene in 180° range, Kinect (vision sensor) showing depth image/RGB image and edge detector output, sonars octagon reading distance in centimeters (proximity sensor - ultrasound), distance access point (AP) to the

robot (simulated received signal strength (RSS) from a wireless network), and robot orientation (yaw).

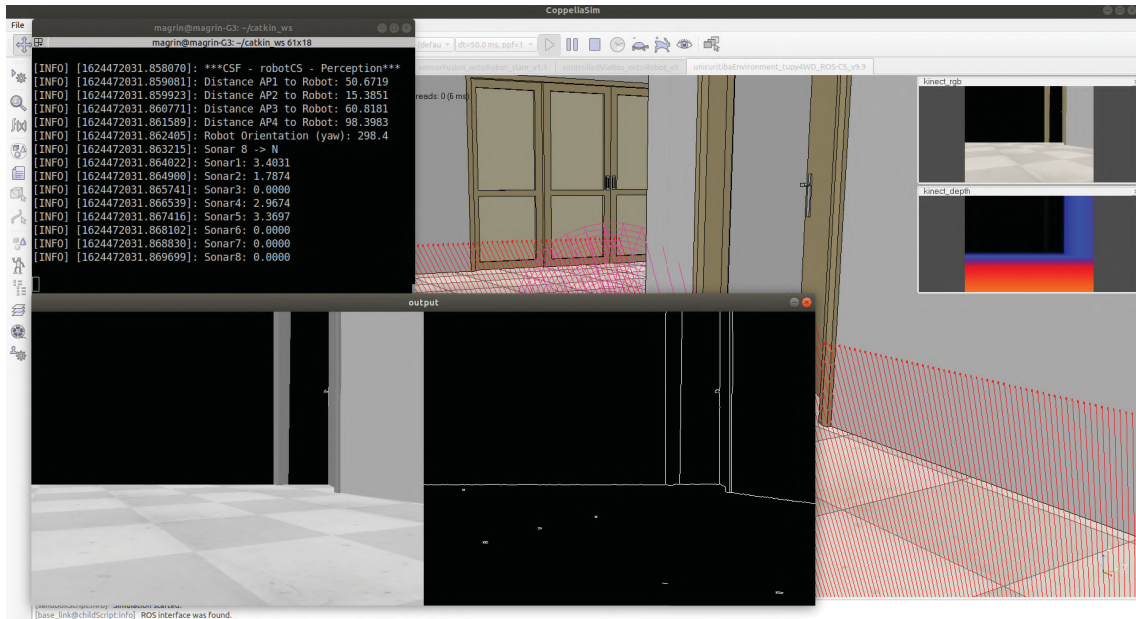


Figura 5.9: CoppeliaSim simulator with ROS interface running ROS packages for the perception of the TUPY-4WD model in the Pinheirinho environment. The figure shows laser scanning the scene, Kinect showing depth image/RGB image and edge detector output, ultrasound octagon reading distance, access point to the robot distance, and robot orientation.

5.1.5 Classification and Object Recognition - NCLT Dataset

To assert the hybrid indoor and outdoor localization and mapping, an essential feature of this proposal, as it is not common in the literature to use both environments, is important to this proposal. We utilize the University of Michigan North Campus Long-Term Vision and LiDAR Dataset (NCLT Dataset) (Carlevaris-Bianco et al., 2016), as the infrastructure and objects in the Michigan North Campus scene are similar to those of the Pinheirinho Campus at the University of Curitiba.

We are training an image for place classification on the NCLT dataset, indoor or outdoor environment. For training, the NCLT dataset of images was organized. The session 2013-04-05 Cam5 started by converting the images to JPEG and manually separating the training base into indoor and outdoor images. To reduce the training time of the model, the images were reduced by approximately one-third, totaling 1908 indoor images and 7890 outdoor images, and were randomly excluded. Figure 5.10 shows the image samples from the indoor and outdoor environments for classification. Experiments with a deep learning algorithm for classification using Keras framework and code example *image classification from scratch* (Fchollet, 2021) to build and train the model. We got to 99% validation accuracy after training for 30 epochs and validated the deep learning to classify environments in a hierarchical cognitive system.

Another experimental method used to identify environments with the adjusted NCLT dataset training classification was Detectron2 object recognition. Facebook AI Research built Detectron2 to support the implementation of object detection algorithms in the PyTorch framework. Figure 5.11 shows the object recognition in an indoor environment to visualize a person, backpack, chair, and table, and in an outdoor environment, visualize a car and a person. There are many



Figura 5.10: Sample images from the NCLT dataset, session 2013-04-05. The figure shows images from places, such as research labs, central aisles, streets, and sidewalks (Carlevaris-Bianco et al., 2016).

approaches to place recognition, and using object recognition is a different way to distinguish between indoor and outdoor in dynamic environments.

The initial experiments with the TUPY-4WD robot and simulations conducted in the Gazebo and CoppeliaSim environments were fundamental for gaining the first insights into sensor integration, perception, and robot locomotion in hybrid environments. Although this project was not directly incorporated into the final version of the mobile platform, it served as an essential foundation for validating the proposed concepts, identifying practical challenges, and guiding the system's development. Thus, the results and lessons learned during this preparatory phase significantly contributed to the consolidation of subsequent project stages, ensuring greater robustness and effectiveness in the development of the VRI4WD platform.

5.2 VRI4WD MULTI-SENSOR MOBILE ROBOT

We built the VRI4WD robot, focusing on studying the main types of low-cost sensor arrangements used in mobile robotics, considering locomotion in a hybrid environment. The datasets available in the literature (Geiger et al., 2013; Maddern et al., 2017; Qingqing et al., 2022) work with perception directed to the type of mapping environment, usually based on high-end and expensive sensors. The proposal in this project is to present a dataset based on sensor blend sets, built on arrangements of low-cost sensors, appropriately representing both indoor and outdoor conditions. This work used the VRI4WD mobile robot platform in the perception-action cycle in the cognitive blended sensors concept. The choice of sensors for sensory hierarchy in the exteroceptive and proprioceptive levels combined the main sensors used in mobile robotics identified in the systematic mapping study (SMS) on multi-sensor fusion (Magrin et al., 2019).



Figura 5.11: Running the model in the Detectron2 library, images show object recognition in the samples of the NCLT dataset. The figure shows object recognition, such as a person, a backpack, a chair, a table, and a car.

The proposed work assists sensor fusion for wheeled autonomous mobile robots using a model of sensor blend sets (SBS). A survey of the main datasets using multi-sensors shows that the platforms use the best sensors found on the market to build the dataset, but do not use different types of sensors highlighted in the SMS (Magrin et al., 2019), see Table 3.1. These sensors were selected for the sensor blended set model. Besides the limited variety of sensors in the available datasets, the platforms employed are not genuine mobile robot models but adaptations of conventional vehicles such as cars, pickup trucks, and a Segway Dicycle, used for locomotion during sensor mapping. The development of a mobile robotic platform dedicated to applying multi-sensors in a hybrid environment is presented as a tool for the SBS model.

The VRI4WD mobile robot is an indoor and outdoor platform with proprioceptive and exteroceptive multi-sensors. Figure 5.12 shows the sensor suite on the VRI4WD, such as GPS NEO-M8M, Laser Scanner RPLIDAR A1, IMU (Accelerometer, Digital Compass, and Gyroscope) BNO055, Ultrasound URM37, Vision Xbox 360 Kinect, Infrared IR Ranger Sharp, and Quadrature Encoder. The VRI4WD is an *open-source* project, available in <https://github.com/VRI-UFPR/vri4wd>.

Figure 5.13 shows the VRI4WD Robot hardware diagram. The *low-level* microcontrollers, ATmega2560 and ATmega328p, read and preprocess sensor data, control the motors and communicate with the *high-level* microcontrollers in the Raspberry Pi boards (“LIDAR” and “KINECT”), using a logic level converter (5V-3.3V). Sensors were identified by colors representing SBS with different features, such as blue blocks (state) with bumper, battery current, and gyroscope information; green blocks (movement) with information from the accelerometer and quadrature encoders; yellow blocks (distance) with information from three ultrasounds, digital compass, and three infrared sensors; orange blocks (mapping) with laser scanner and vision information; red blocks (positioning) with information from the GNSS (GPS, Galileo,

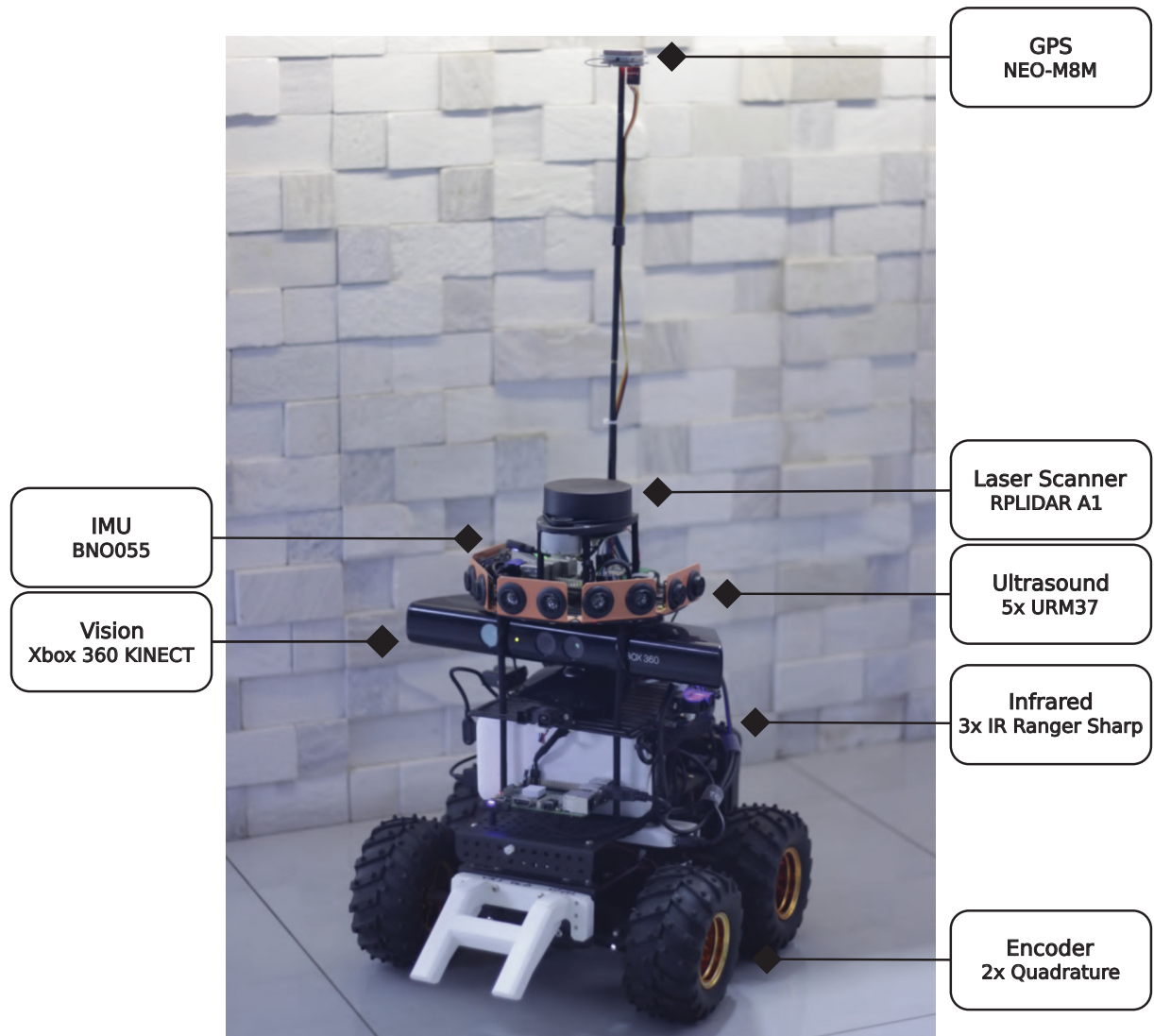


Figura 5.12: Sensors suite on the VRI4WD Mobile Robot Platform. The figure shows multi-sensors in the robot, such as laser range scanner RPLIDAR A1, ultrasonic distance sensor DFRobot URM37, IMU with absolute orientation sensor Bosch BNO055, camera Microsoft Kinect XBOX 360, infrared DFRobot Sharp GP2Y0A02YK, quadrature encoders embedded in the front motors, and GPS u-blox M8 GNSS fixed on the top.

GLONASS, and BeiDou) and RF sensor. The hardware block diagram also organizes the power supply modules, with a 5V/5A power module from DF Robot, a DC-DC converter (Buck-Boost) to power the Kinect at 12V, and a voltmeter to monitor the LiPo battery. Power banks helped power the Raspberry Pi. H-Bridge Module, DC Motor Drive, provides current and control to the robot's four motors powered by the LiPo battery. Table 5.1 shows the VRI4WD sensors organized by the five SBS levels, sensor measurement type, proprioceptive or exteroceptive, number of sensors, and the part number for reference.

The mechanical platform used a 4WD (Four Wheel Drive) all-terrain chassis from Dagu, the Wild Thumper model. The chassis has independent suspension for each of its spiked 120mm wheels (Figure 5.14).

The Robot Operating System (ROS) provides services designed for hardware abstraction. ROS connection information from the VRI4WD robot sensors to *nodes* so that they can transmit messages among them through topics to *publish* and *subscribe* data. Figure 5.13 shows the

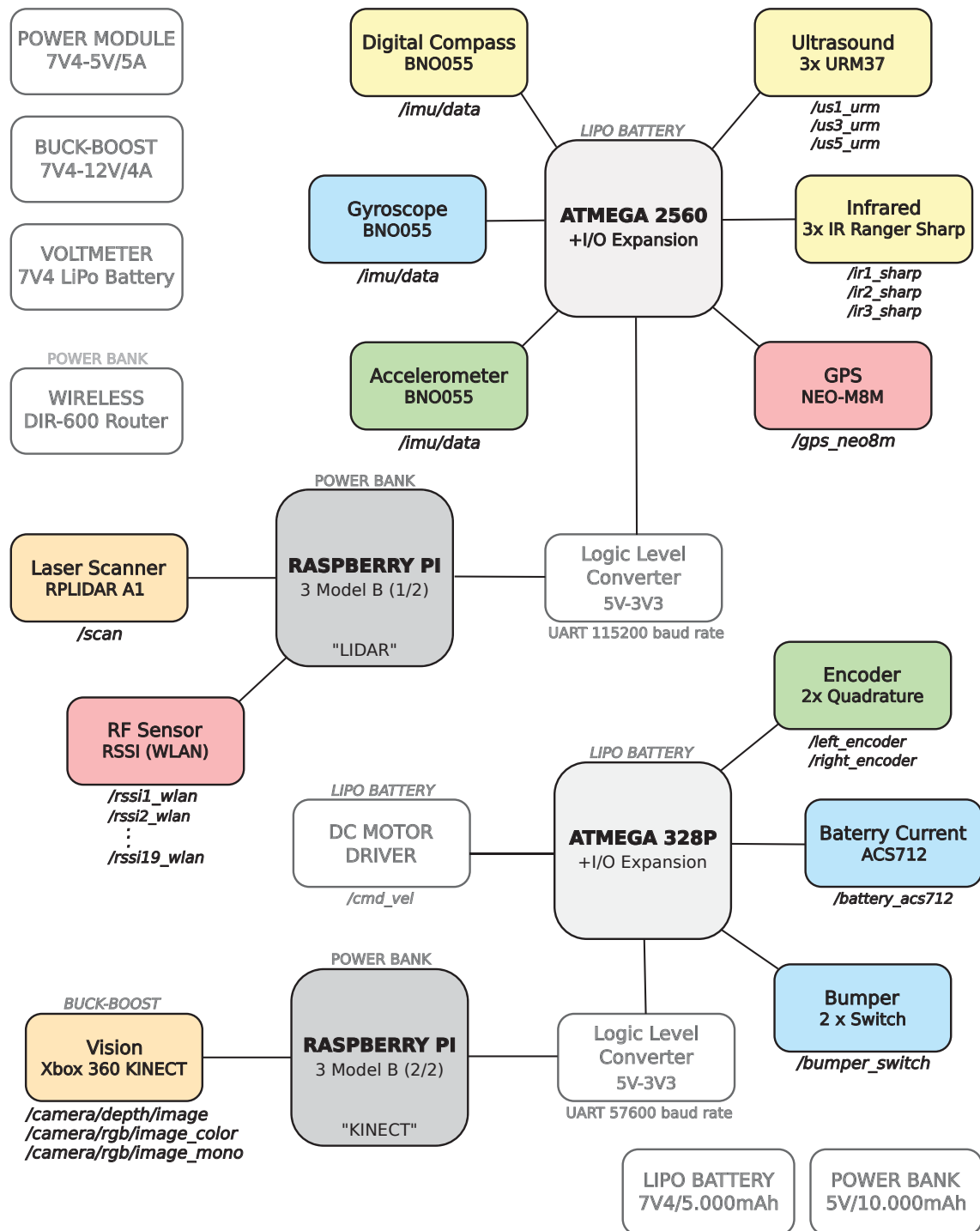


Figura 5.13: VRI4WD Robot hardware block diagram. At the top - are three blocks, and at the bottom - are two blocks representing the DC-regulated power for supplying the motors, electronic boards, and sensors for each voltage level. Light gray blocks represent the 8-bit microcontrollers, input analog and digital sensor data, DC motor driver output controls, and UART connection with the Raspberry Pi boards using a logic-level converter. Sensors were identified by colors representing SBS with different features.

subscriber topics of each type of sensor and *publisher* topics for DC motor control. The Wi-Fi router is embedded in the platform for easier communication via ROS, with a notebook (Fig. 5.15).

Presentation of ROS *topics* related to each microcontroller system (Figure 5.13), such as:

ATmega 328P 8-bit 32KB 23GPIO

Tabela 5.1: The VRI4WD robot sensors. The table was divided into five levels of the SBS model, relating the type of measurement proprioceptive or exteroceptive, the type of sensor, and the part number.

SBS	Measure	Sensor	Model
State	Proprioceptive	Battery current	ACS712
	Proprioceptive	Gyroscope	BNO055
	Exteroceptive	Bumper	Contact switch
Movement	Proprioceptive	Accelerometer	BNO055
	Proprioceptive	Quadrature encoders	Motor integrated 48 CPR
Distance	Exteroceptive	3x Ultrasound	URM37 (range 2 ~ 800cm)
	Exteroceptive	Digital Compass	BNO055
	Exteroceptive	3x Infrared	GP2Y0A02YK (range 20 ~ 150cm)
Mapping	Exteroceptive	Laser scanner	RPLIDAR A1 (range 12m, 360°)
	Exteroceptive	Vision	Xbox 360 Kinect
Positioning	Exteroceptive	GPS	NEO-M8M (up to 3 GNSS)
	Exteroceptive	RF sensor	WLAN - RPi "LIDAR"



Figura 5.14: VRI4WD Mobile Robot Platform. The mechanical platform is a 4WD all-terrain chassis. (a) The Figure shows the robot moving along the trajectory in an indoor environment. (b) The Figure shows the robot moving along the trajectory in an outdoor environment.

- DC Motor | *subscriber /cmd_vel* - subscribes the linear and angular velocity in meters per second of the robot (*Message type geometry_msgs/Twist*).

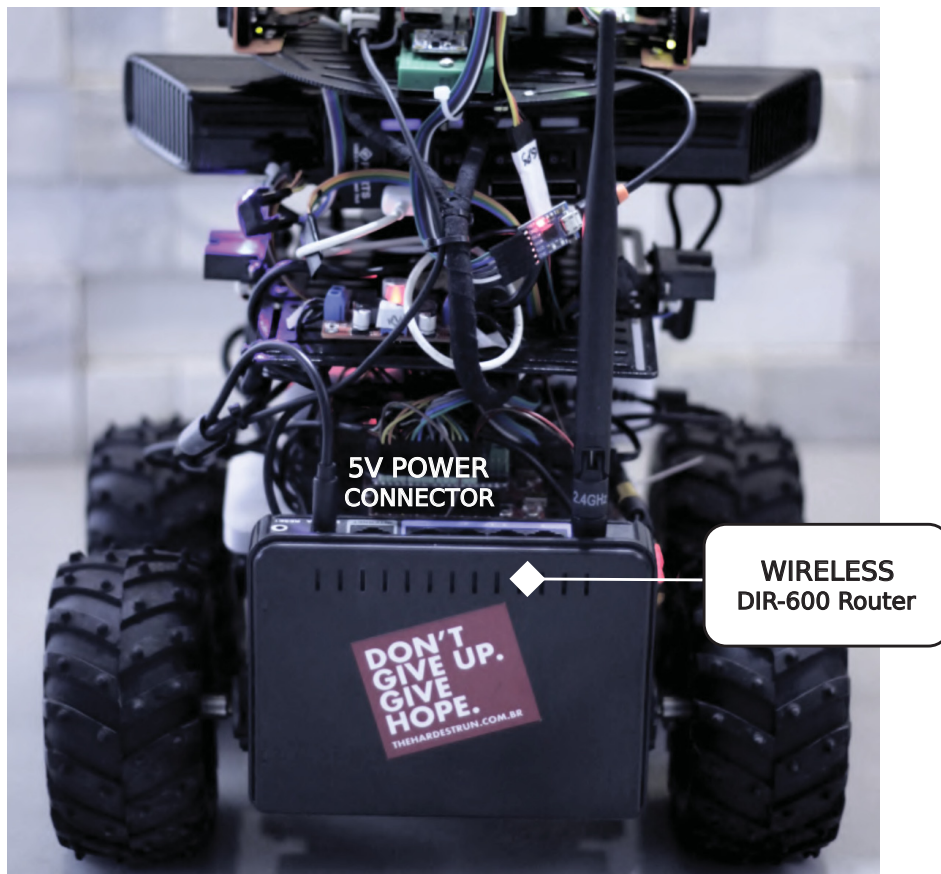


Figura 5.15: VRI4WD robot with Wireless D-Link DIR-600 2.4GHz IEEE 802.11b/g router embedded on the rear of the platform for more accessible communication with notebooks via ROS. The wireless router has a 5V DC / 1A power connector and an external fixed 5dBi antenna.

- Encoder | *publisher /left_encoder, /right_encoder* - publishes left encoder and right encoder data in pulses (*Message type std_msgs/Int16*).

ATmega 2560 8-bit 256KB 86GPIO

- Accelerometer | *publisher /imu/data* - publishes the linear acceleration in meters per second (*Message type sensor_msgs/Imu*).
- Digital Compass | *publisher /imu/data* - publishes the orientation in quaternion form (*Message type sensor_msgs/Imu*).
- GNSS | *publisher /gps_neo8m* - publishes the latitude and longitude degrees using the WGS 84 reference (*Message type sensor_msgs/NavSatFix*).
- Gyroscope | *publisher /imu/data* - publishes angular velocity in rads per second (*Message type sensor_msgs/Imu*).
- Infrared | *publisher /ir1_sharp, /ir2_sharp, /ir3_sharp* - publishes the range of the front infrared, left infrared and right infrared in centimeters (*Message type sensor_msgs/Range*).
- Ultrasound | *publisher /us1_urm, /us3_urm, /us5_urm* - publishes the range of the front ultrasound, left ultrasound and right ultrasound in centimeters (*Message type sensor_msgs/Range*).

Raspberry Pi 3 Model B “KINECT” 64-bit quad-core ARM Cortex-A53 1GB

- Vision | *publisher /camera/depth/image, /camera/rgb/image_color, /camera/rgb/image_mono* - publishes depth, color and monochrome images with a resolution of 640 x 480 pixels, at 30 fps (*Message type sensor_msgs/Image*).

Raspberry Pi 3 Model B “LIDAR” 64-bit quad-core ARM Cortex-A53 1GB

- Laser Scanner | *publisher /scan* - publishes scan topic from the laser (*Message type sensor_msgs/LaserScan*).
- RF Sensor | *publisher /rssi1_wlan, /rssi2_wlan, /rssi3_wlan, /rssi4_wlan, /rssi5_wlan, /rssi6_wlan, /rssi7_wlan, /rssi8_wlan, /rssi9_wlan, /rssi10_wlan, /rssi11_wlan, /rssi12_wlan, /rssi13_wlan, /rssi14_wlan, /rssi15_wlan, /rssi16_wlan, /rssi17_wlan, /rssi18_wlan, /rssi19_wlan* - publishes the received signal strength indication (RSSI) value module data in dBm from each of nineteen addressed access points (*Message type sensor_msgs/Int8*).

The VRI4WD platform, Figure 5.16, was assembled to work with many sensors and enable adjustments and modifications in validating different types of sensors. The onboard electronics were segmented as *low-level* boards (8-bit ATmega microcontrollers) receiving digital or analog information from sensors or controlling the motors. *High-level* electronic board (Raspberry Pi 3B), processing LiDAR sensor or Kinect packages and providing a Wi-Fi interface for the RF sensor.

Figure 5.17 shows the VRI4WD Robot sensors on the platform with the LiDAR at the top for a 360-degree reading. The sonars are positioned just below for the same perception of objects in the scene. We set empirically the Kinect camera as high as possible for a better viewing angle, but without disturbing the reading of the other sensors. The infrared sensors measure the distance in centimeters to an object and are positioned to the robot’s front, right, and left (Figure 5.18). Figure 5.19 shows the IMU module (BNO055) was positioned on the top level of the platform and close to the center of the VRI4WD robot. The assembly on a breadboard seeks to make the exchange between modules of different models more flexible for validation when applying multi-sensors in a hybrid environment. We used a USB-UART converter module to communicate with the Atmega 2560 microcontroller.

There were numerous difficulties in building and validating the VRI4WD platform. With the proposal to collaborate in the construction of new robots, the main challenges and solutions found for validation are listed in the following:

- **Robot overturning:** *challenge* - large number of sensors integrated into the platform, the center of gravity was very high and when a strong movement the robot tipped forward; *solution* - distribute the weight along the platform with the batteries close to the wheels, wireless router positioned at the rear and to ensure the robot does not tip over on rougher outdoor terrain, a front bumper was developed and manufactured in the ROSIE laboratory at UniCuritiba for frontal installation on the platform.
- **Power supply:** *challenge* - provide voltage and current levels for motors, electronic boards, and sensors. Initially, a 7.4V/5A LiPo battery, 5V regulator module, and boost-type DC-DC converter set at 12V were used to supply power to the entire robot. The current provided by the battery was serving the robot. Still, there were some instabilities in the operation of the Raspberry Pi, causing the OS to freeze and not recognize the laser scanner connected to the USB port due to the instability in the

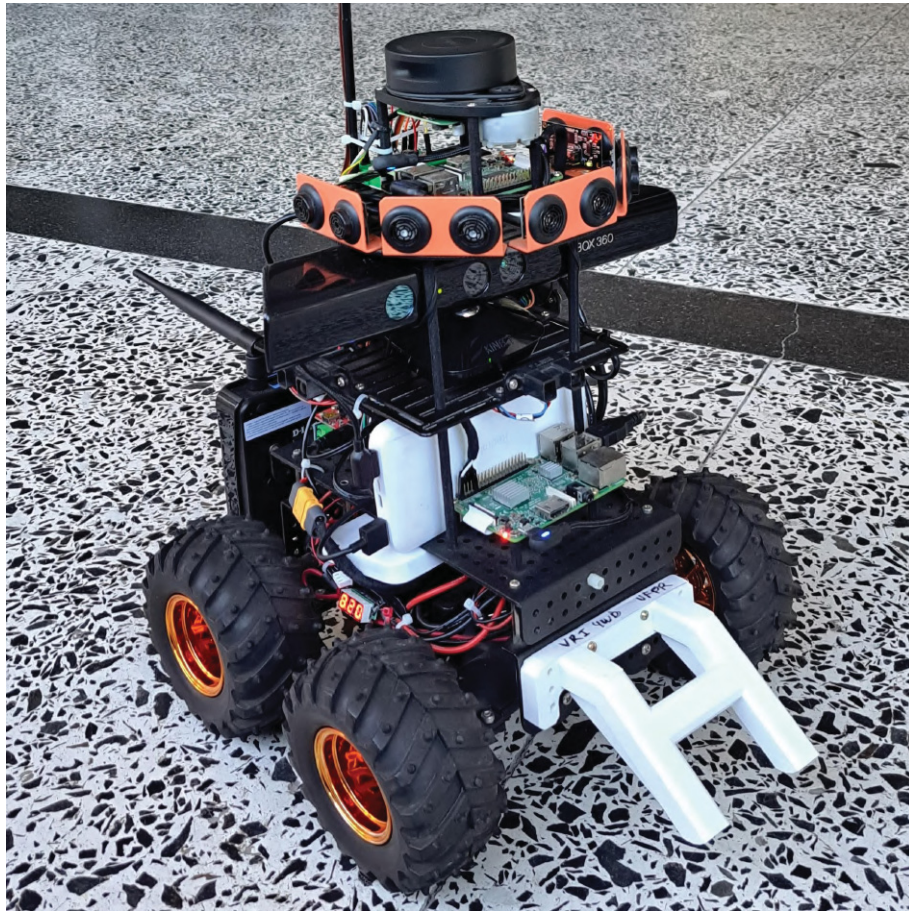


Figura 5.16: The position of the sensors at different levels ensures better recognition of the environment, ease of adjustment, and better validation of the other sensors. We balanced the robot's center of gravity with the position of the batteries at the bottom and the router at the rear. A bumper was mounted on the front of the robot to prevent it from tipping over while moving on uneven terrain.

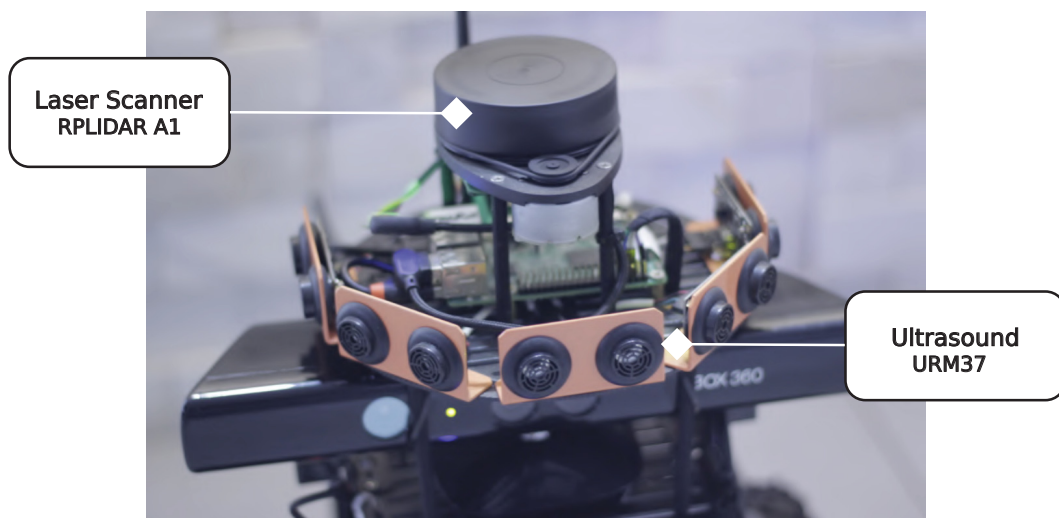


Figura 5.17: Sensors are positioned at different levels on the VRI4WD robot platform. At the top, the laser scanner (RPLIDAR A1) and five ultrasound sensors.

voltage level; *solution* - segment the power supply into different sources, leaving the 7.4V/5A LiPo battery for the motors, ATmega boards, and Kinect. Raspberry Pi boards

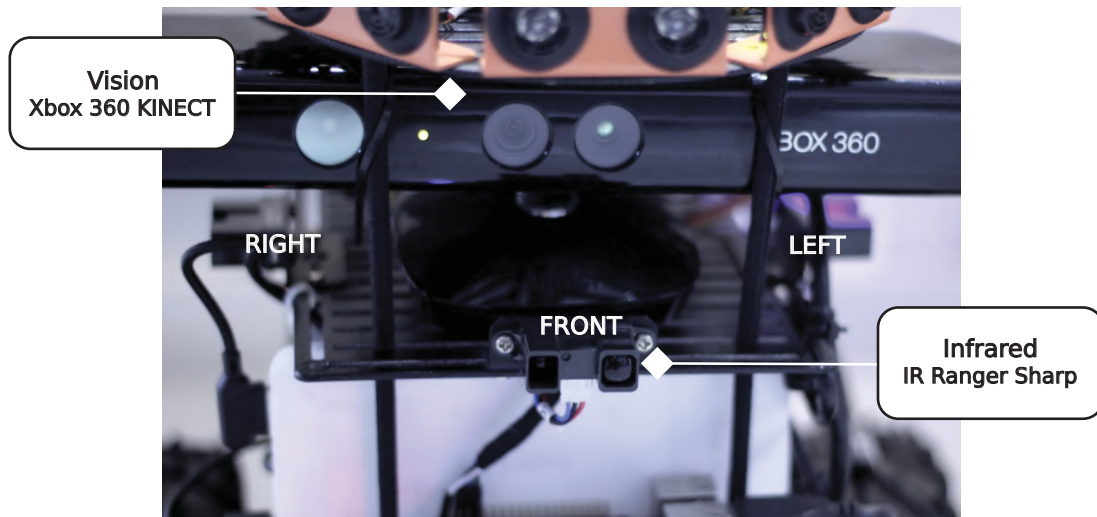


Figura 5.18: The infrared sensors are at the same level as the Kinect and positioned front, left, and right to measure the distance in centimeters to an object.

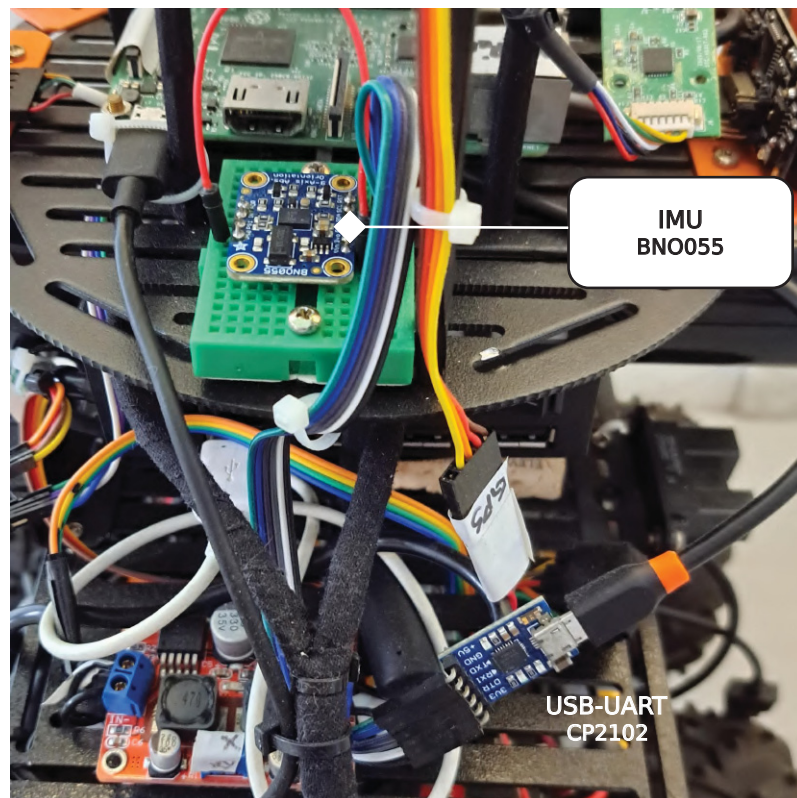


Figura 5.19: The IMU BNO055 was on the last platform close to the center of the robot. The module was on a breadboard, as it was necessary to test different models. The BNO055 uses UART communication, converter module USB-UART (CP2102).

worked well using a dedicated 5V/10A power bank, and another 5V/10A power bank was used for the wireless router.

- **Inertial measurement unit:** *challenge* - reading information from the accelerometer and gyroscope using the MPU-6050 single-chip 6DOF IMU and GY-85 9DOF IMU module with the integrated HMC5883L - digital compass (magnetometer), ITG3205 - gyroscope and ADXL345 - accelerometer. Despite the ease of reading through I2C

communication, both solutions presented noise, causing uncertainty in measurements; *solution* - using the proposal of a Bosch Sensortec BNO055 IMU with integrated sensor fusion and delivering the output data in quaternion and Euler angles for applications in robotics using ROS.

- **GNSS interference:** *challenge* - influence when turning on Kinect on the GPS NEO-5M module positioning update. GPS stops receiving the positioning of satellites, and only by turning off the Kinect does the GPS reading return; *solution* - move the GPS module away from the magnetic interference generated by the Kinect. The module was positioned with spacers at a height of 45 cm from the Kinect and replaced by a GNSS NEO-M8M multiple positioning module (GPS, Galileo, GLONASS, BeiDou). Acquisition of internal and external positioning works even with Kinect turned on.
- **Controlling the robot:** *challenge* - control the robot for locomotion while acquiring mapping data in internal and external environments. A Turnigy 9X radio receiver was installed to control the robot via the ATmega328p board, but it presented interference with the wireless communication of the Raspberry Pi, which operates via the same 2.4GHz RF; *solution* - firstly, it was to exclude the robot's Turnigy receiver which was generating interference when the transmitter was connected and control the robot via the keyboard using a ROS topic. However, using the keyboard during mapping hindered the process, which required attention when recording data at each landmark. Therefore, an XBOX USB joystick was added using the *teleop_twist_joy* package to control the robot's locomotion during mapping.
- **Fixed suspension:** *challenge* - as the high center of gravity combined with the numerous onboard sensors caused the platform to sway excessively during locomotion. This instability negatively impacted image acquisition and the smooth operation required for accurate encoder data collection. In irregular terrain, instead of providing support, the suspension occasionally caused the wheels to spread and drag along the ground, further compromising the sensors' mapping; *solution* - dedicated supports were developed to fix the suspension in place, ensuring more controlled platform movement and improving the overall quality of the sensor data (Figure 5.20).

Another area for improvement was with the USB cables, both from the Laser Scanner, which, in addition to the data, is powered by the cable itself, and from the USB to DC 5.5 mm adapter of the wireless router. These cables are causing instability in the functioning or even recognition of devices. Changing to better quality cables solved the problem, but it is a problem that, if not initially detected, could take a long time to validate the project.

In addition to the challenges encountered in developing the VRI4WD robot due to the large number of low-cost sensors on a compact platform for indoor and outdoor use, it was necessary to create a dedicated Raspberry Pi module for recording the ROSBAG. During dataset validation, we observed encoder pulses with count discontinuities and delays in Kinect images due to wireless communication, as data collection occurred on the same Raspberry Pi as the camera (Kinect). For a new mapping session, the wired Raspberry Pi modules were placed on the robot-embedded router, modifying the Kinect settings to reduce the number of frames without compromising the dataset's proposed parameters. We captured the dataset using a custom data acquisition system (DAS) on the mobile robot, Raspberry Pi 3 B+ (ROS Melodic in Ubuntu 18.04 LTS).

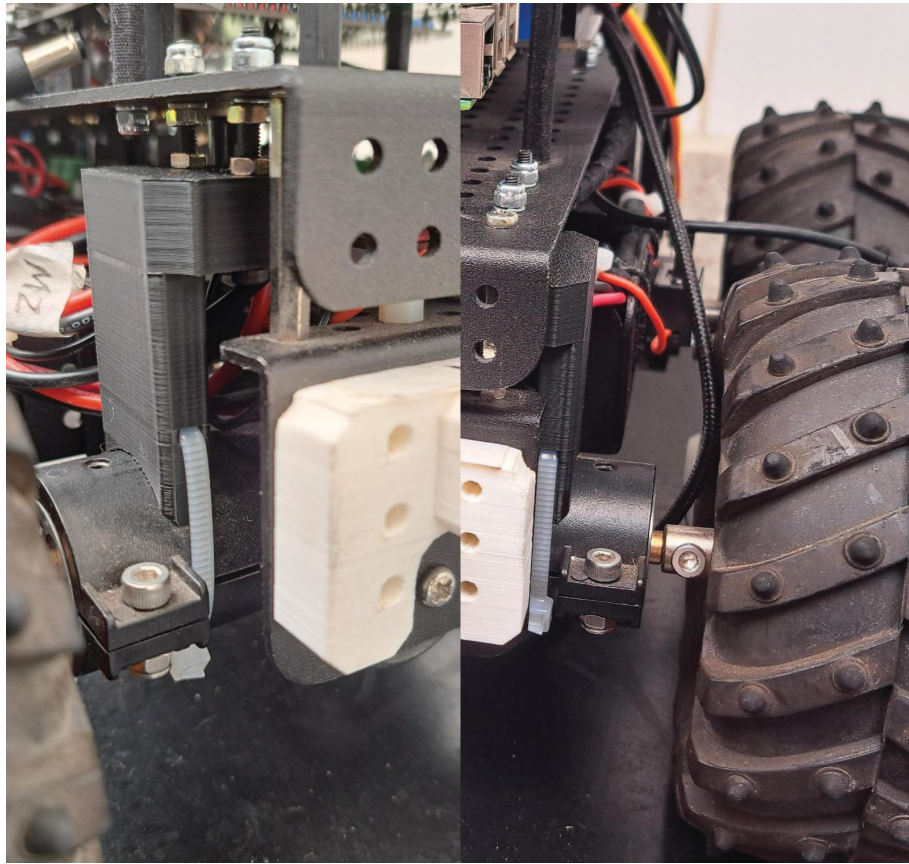


Figura 5.20: An adapter is needed to attach the VRI mobile robot's 4WD suspension. The figures show the adapter with height adjustment using upper screws and a clamp for attachment to the motor shaft, locking the suspension and preventing unnecessary platform swinging.

5.3 CHAPTER CONSIDERATIONS

This chapter presents the experimental guidance for developing the VRI4WD mobile robotic platform. Early experiments, utilizing simulation environments and existing datasets, provided foundational and validated initial concepts. To complement this work, the VRI4WD mobile robot, equipped with multi-sensors, was developed aimed at locomotion in hybrid environments. Custom-built with several low-cost sensors and designed for robust indoor-outdoor navigation, the VRI4WD addresses a significant gap in available research platforms. Its open-source design for adaptability makes it a valuable tool for future studies in multi-sensor fusion and robotic perception.

6 BUILDING A DATASET

The process of constructing the UFPR-MAP dataset, including the methodology for collecting a geodetic ground-truth with transitions between indoor and outdoor environments, is presented. This dataset has unique characteristics, such as integrating multiple types of low-cost sensors and geodetic ground-truth in a hybrid environment, representing a fundamental contribution to this research.

6.1 BUILDING THE VRI4WD UFPR-MAP DATASET

The VRI4WD UFPR-MAP dataset is a valuable resource for researching the blend of sensors in indoor and outdoor environments. This work contributes to the mobile robot's research community by making available a dataset with coordinates of points obtained by geodetic and topographic methods, enriched with several sensors' data.

6.1.1 Creating a ground-truth

To validate the mapping of sensors in the environment, it was necessary to position some landmark points on the map to create accurate position information to ground-truth the robot trajectory. The densification of the positioning was initially defined regarding the precision necessary for the dataset proposal and adjusted for the building's architecture, considering the average distance between the building's pillars of 3 meters. The mapping was carried out with a pair of GNSS receivers and a total station, following the parameters:

- GNSS Receivers: Topcon Hiper Lite Plus (5mm + 5ppm (horizontal precision) and 20mm + 1ppm (vertical precision));
- Reference System: SIRGAS 2000;
- Projection System: UTM (Universal Transverse Mercator, CM 51W, Zone 22);
- Total Station: Leica TCR 407 (7" angular precision and 2mm + 2ppm (linear precision));
- Densification: 3 m (The distance between the pillars of the building is used as a standard);
- Environment: hall (indoor) and sidewalk (outdoor);
- Marking: numeric tag (VRI4WD UFPR-MAP).

Professor Wander da Cruz conducted the measurements from the Geodetic Instrumentation Laboratory (LAIG) at the Federal University of Paraná (UFPR). First, a point was implanted, and then the coordinates of an already established point were determined. These coordinates were later used as a reference and guidance in the survey with the total station.

The static relative method was applied to determine the coordinates, which occurs when the Cartesian coordinates of an unknown station are chosen from a known station, with the main observable used in this type of positioning being the phase of the carrier wave (LEICK, 2004). Figure 6.1 represents the relative positioning, with station "A" having known coordinates and station "B" determining. Topcon receivers with an occupancy time of 2 hours were used to

determine the coordinates. After the tracking time was completed, data processing was carried out using the TopconTools software, using the Brazilian Network for Continuous Monitoring of Systems (*Rede Brasileira de Monitoramento Contínuo dos Sistemas (RBMS)*) station of the GNSS systems as a base station for applying the relative static method. Figure 6.2 shows the GNSS receiver over one of the orientation points and the baselines generated in the static relative processing in the TopconTools with the distances to the base station.

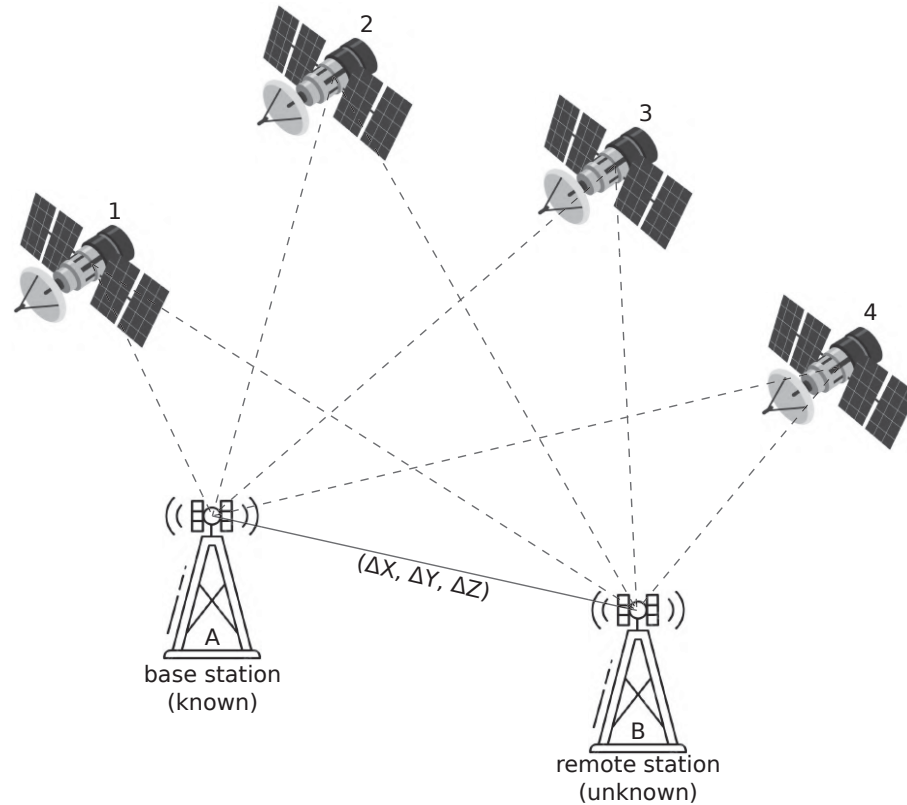


Figura 6.1: Relative Method. The coordinates of the unknown point “B” are calculated using the known coordinates of point “A” as a base. Adapted from (SUCI, 2012).

To carry out the topographic survey with the total station, the coordinate values obtained by GNSS were used, processed in the SIRGAS 2000 reference system, and converted to the UTM Projection system in the TopconTools processing software itself. With a tracking time of 2 hours and short baselines, with a maximum of 571.910 m (RN2053D), the horizontal accuracy of the reference points was in the order of 0.003 m.

The framed polygonal method was applied to the topographic survey. According to Veiga et al. (2007), in this method, we start from two points with known coordinates and end at two other points with available coordinates (Figure 6.3), making it possible to calculate the azimuth and the distance between the arrival points and compare them to the values obtained from field measurements, enabling the control and distribution of linear and angular errors.

In this work, we used the two points with known coordinates, RN2053D and RE01, at departure and arrival, as shown in Figure 6.4. The points were obtained through topographic irradiation, with coordinates calculated from the support polygonal and used to validate the coordinates obtained by the other sensors.

The polygon had an angular closure error of 27” and a linear error of 1/42.910. This angular error means 4.5” distributed among the six occupied points. The linear error corresponds to 1 millimeter for every 42.9 meters traveled, representing an error close to 7 mm for the entire polygon. This value meets the system requirements to be validated.

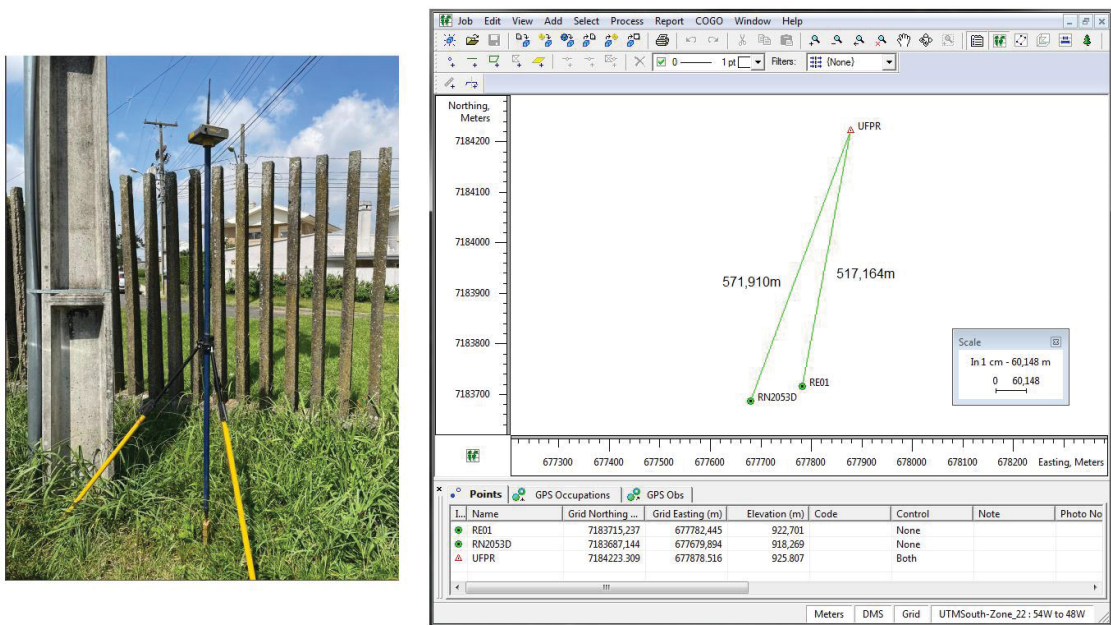


Figura 6.2: GNSS and baselines. GNSS receiver over point RE01 was installed on a 2.00 m antenna, and an image from the processing software showed the vectors that were part of the baselines.

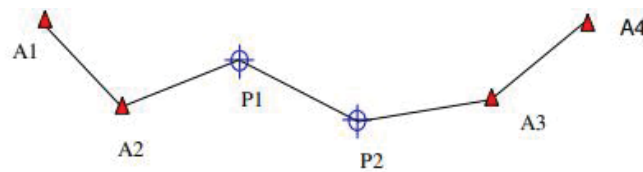


Figura 6.3: Example of a framed polygon, starting A1 and A2 points and ending A3 and A4 points with known coordinates. P1 and P2 represent the pole (prism) to reference the positioning (Veiga et al., 2007).

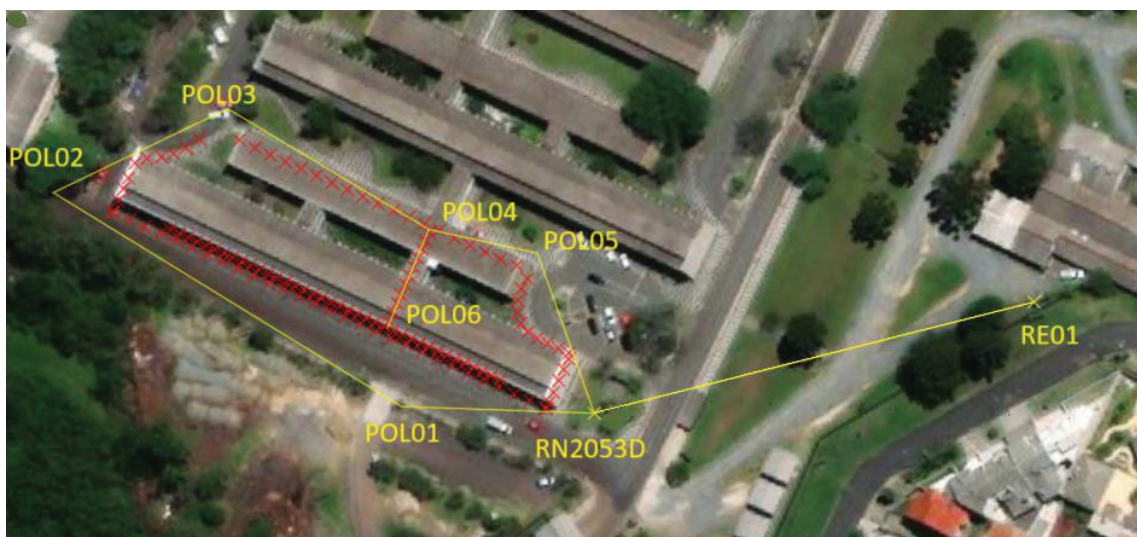


Figura 6.4: Support polygonal and points for validation. The support polygon, along with the names of the points, is presented in yellow, and the points used to validate this work are represented in red.

Figure 6.5 shows the measurement procedure with the total station, using the pole (prism) to reference the landmark points and tie the outdoor positioning with the indoor one.

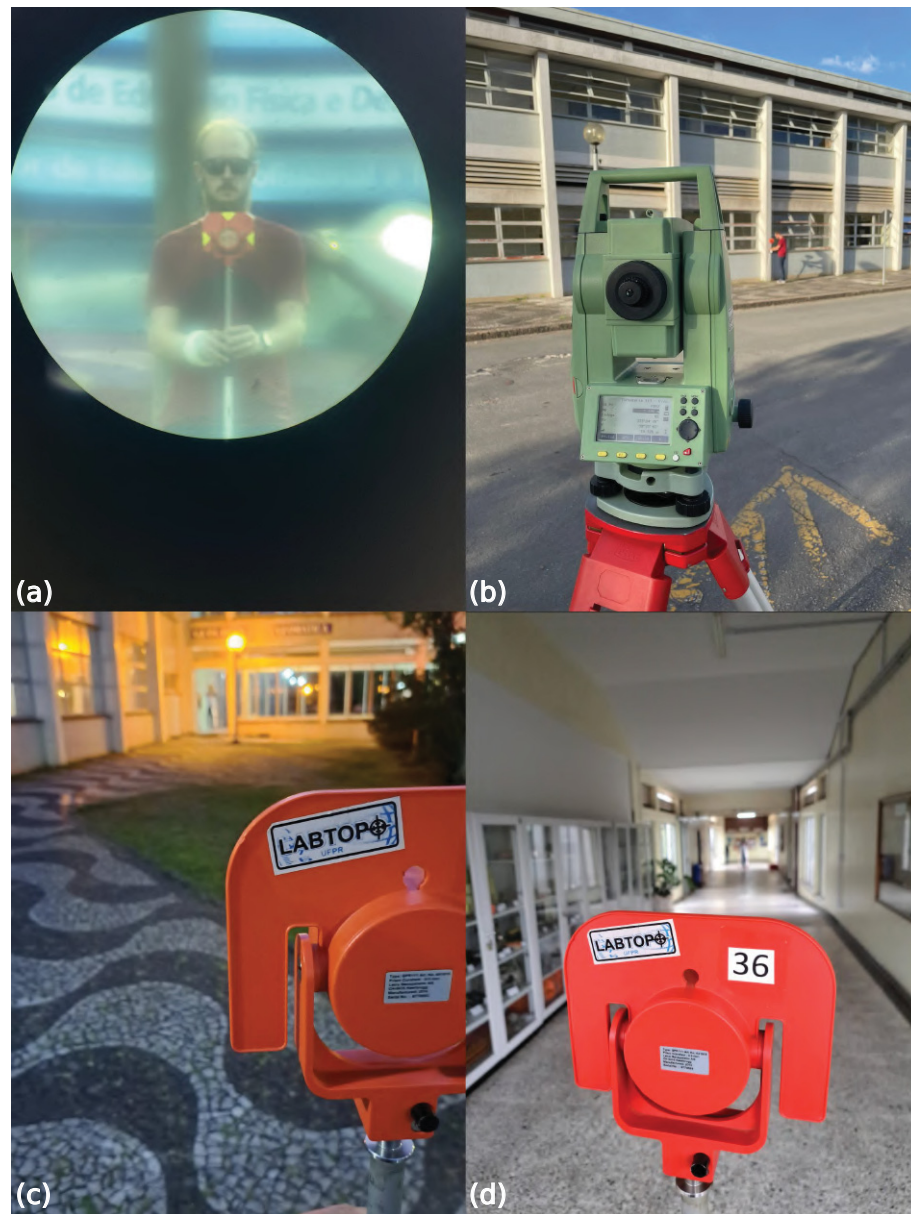


Figura 6.5: Measurement of each point in the building of the Geology Department in Engineering Street from Polytechnic Center - UFPR. (a) The procedure with the total station, using the pole (prism) to reference the landmark points. (b) Total station Leica TCR 407 from LAIG/LABTOP UFPR for the measurement of each point. (c) The figure shows the procedure for tying the outdoor points with the indoor ones made with the total station. (d) The method for tying the indoor points at the building.

The trajectory of the VRI4WD robot in the Geology building, Engineering VI, was represented on the Google Earth map, following clockwise from the aisle inside the building to the sidewalk outside the campus to return inside. Figure 6.6 shows the position numbers and the points of each trajectory alphabetically, robot locomotion on a total path (indoor and outdoor) of approximately 350 meters.

The mapping used tags to reference each point on the map, numbering each position. We used the colors gray and white in multiples of five to differentiate the distance between the tags; that is, gray tags with an approximate distance of 3 meters and white tags with an approximate distance of 15 meters. Points represented by white tags were used as previous readings of access point signals presented at the building. Figure 6.7 shows the tags manufactured using a printer liquid resin, indicating the name of the dataset and its position number on the map.



Figura 6.6: VRI4WD UFPR-MAP, positions in a Google Earth map. The numbers indicate the 123 landmark points on the Polytechnic Center - Geology Building. It alphabetically represents the point-to-point trajectory of the robot on the map, indoor (AB and BF), and outdoor (BC, CD, DE, and EB).



Figura 6.7: Tags manufacturing in a printer, liquid resin. Showing the name of the dataset and the position number on the map. Tags with a diameter of 30 mm and a thickness of 2 mm, using a projection for the letters and a recess for the number. Used in two colors, gray or white, to differentiate multiples of 5 from white tags.

The tags were fixed to the building wall at an approximate height of 200 mm from the ground, and to represent the position of the trajectory on the sidewalk, some pillars were used as reference points. Figure 6.8 shows the position of the tags in the outdoor and indoor environments of the Geology building. One hundred twenty-three tags were positioned around the building, approximately 3 meters from each other.



Figura 6.8: Tags fixed to the wall and pillars around the Geology building at the Federal University of Paraná - Polytechnic Center. The tags were positioned at an approximate height of 200 mm from the floor and tried to follow the architecture of the building to place the tags on each column or pillar, representing an approximate distance of 3 meters between the tags.

6.1.2 UFPR WLAN map

Taking advantage of the wireless network structure for data collection from available access points (APs) in the environment has been a strategy adopted in many works about wireless indoor localization (Kunemund et al., 2009; Carvalho et al., 2012; Yang and Zhang, 2014; Torteeka and Chundi, 2014; Magrin and Todt, 2016). This work used wireless signal (RSSI) from APs, measuring each registered MAC address in dBm.

Before using the Wi-Fi network as a reference, all APs in the geology building were mapped out. WLAN readings were carried out at multiple points of five (white tags), so the approximate distance between them is 15 meters, and a smaller distance would not result in a better correspondence between the collected points. APs were selected based on their appearance in at least 6 points on the map and with signal quality more significant than 40% to avoid the unavailability of points in the collected regions. Table 6.1 shows the nineteen WLANs selected as readings from the RF sensors of the VRI4WD robot. Each RSSI was registered with its MAC address.

Tabela 6.1: UFPR WLAN MAP. Nineteen RSSI were selected as RF sensors of the VRI4WD robot.

RSSI	Name	Address
1	VRI4WD_WLAN	1C:AF:F7:9B:CC:04
2	UFPR_SEM_FIO	20:BB:C0:F9:58:F0
3	UFPR_SEM_FIO	20:BB:C0:F9:58:FF
4	UFPR_SEM_FIO	50:06:04:2B:D7:50
5	eduroam	20:BB:C0:F9:58:F1
6	LAIG 1	9C:53:22:63:5A:C3
7	eduroam	50:06:04:2B:D7:51
8	SALAMUNI	00:21:27:D0:85:A4
9	LAPE	10:FE:ED:C4:BA:BA
10	SR2015	14:CC:20:9A:14:3A
11	Lab_FOTO	18:A6:F7:38:8F:40
12	LABCARTO	58:10:8C:2F:D8:C6
13	LABAPWan2.4GHz	58:D5:6E:C8:1B:CF
14	Renatin02	64:DB:F7:30:98:69
15	Foyer CJ	84:D8:1B:3C:C5:C6
16	Macmini's Wi-Fi	90:72:40:25:87:6C
17	UFPR_SEM_FIO	50:06:04:2B:DE:90
18	eduroam	50:06:04:2B:DE:91
19	UFPR_SEM_FIO	50:06:04:2B:E0:70

6.1.3 Sensor mapping dataset

One of the contributions of the work is the study of sensor blend sets as a cognitive process of combining multi-sensors. In this way, to meet the most significant number of sensors on a mobile robotic platform, we build the VRI4WD robot, applying the main low-cost sensors used in mobile robotics, that is, standard sensors found in local laboratories, where they can be tested on benches for repeatability of real sensors. Another relevance of applying multi-sensors is exploring a dataset with the main sensors; datasets such as KITTI and NCLT use great sensors, but they are too expensive for an approach with real robots. Sensor mapping using a ground-truth organized by LAIG and taking advantage of the building's WLAN structure complements the relevance of the dataset for studying sensor fusion with attention to the sensor blend sets. This work contributes to a sensor mapping dataset in indoor and outdoor environments. The trajectory environments for the robot, considering different surfaces and architectures along the path (Figure 6.9).

The all-terrain platform of the VRI4WD robot was necessary to explore the entire map when constructing the dataset, so the outdoor surface has some unevenness. For sensor mapping, a joystick, 8BitDo 2.4G Wireless, was used with interaction buttons to register the landmark points, movement controls - angular and linear proportional velocity control using analog sticks, seeking better trajectory control during mapping. Topic analysis was possible through a tablet using the SSH protocol to access the Raspberry Pi (DAS). Before recording, we synchronized all three Raspberry Pi Modules embedded in VRI4WD in time and data. A ROS topic with the value of landmarks is included to facilitate comparison with the ground-truth. Figure 6.10 shows the exploration trajectory environment of the VRI4WD robot controlled by a joystick.

The process for mapping the VRI4WD robot sensors follows the steps: ground-truth UFPR-MAP; VRI4WD robot sensors messages according to ROS topics; access points available



Figura 6.9: The environment used to build the database during the mobile robot's locomotion. Different surfaces and architectures were considered along the path, outdoor environment: (a) similar architecture, long distance for sensor measurement, and linear trajectory; (b) different architectures, parking, and irregular trajectory; (c) proximity between the walls on both sides and transition with the indoor environment; (d) different architectures, parking, and irregular trajectory; indoor environment, (e) narrow aisle for linear trajectory and measurement with left and right side sensors. Figure (f) shows the easy transition between indoor and outdoor environments through two side doors.

at the building (WLAN); record a ROSBag file with the interesting topics; moving the robot through the joystick at a specific point; marking the landmark in each one of the 123 points.

6.1.4 UFPR-MAP dataset

The UFPR-MAP dataset consists of data collected by a VRI4WD Mobile Robot platform (Figure 5.12) on the campus of the Federal University of Paraná. The VRI4WD is outfitted with a GPS NEO-M8M (up to 3 GNSS - GPS, Galileo, GLONASS, BeiDou), laser scanner 360 degree YDLIDAR X2, IMU (Accelerometer, Digital Compass, and Gyroscope) Bosch BNO055, three ultrasound DFRobot URM37, camera Microsoft Xbox 360 Kinect, three infrared ranger Sharp GP2Y0A02YK, two quadrature encoded motors, and RF sensor (WLAN Raspberry Pi). The VRI4WD UFPR-MAP dataset is available in <https://github.com/VRI-UFPR/ufpr-map>.

In order to validate the mapping of sensors in the environment, it was necessary to position some landmark points on the map to create accurate position information to ground-truth

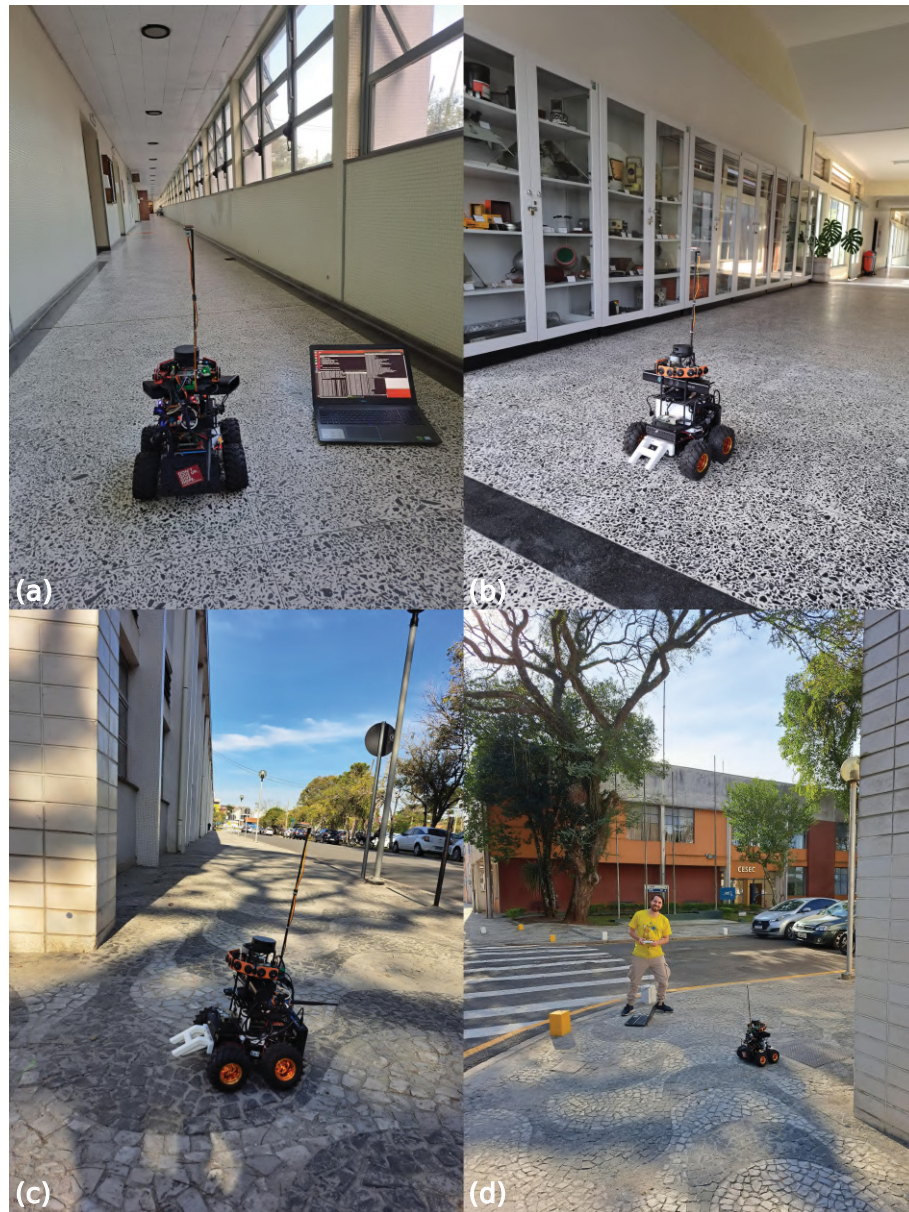


Figura 6.10: VRI4WD robot exploring indoor and outdoor UFPR-MAP. (a) The notebook was used to read the sensor topics through Wi-Fi communication embedded in the robot. (b) A robot is exploring the indoor environment. (c) A robot is exploring the outdoor environment. (d) Move the robot around the map with joystick.

the robot trajectory. To provide ground-truth robot position, we obtained the points through topographic irradiation, a pair of GNSS receivers, Topcon Hiper Lite Plus, and a total station, Leica TCR 407, was used to provide indoor and outdoor points around the building. Points to validate the ground-truth UFPR-MAP are shown in Figure 6.4. The list of ground-truth points was used as a reference landmark for mapping sensors in the environment. The point values were collected in UTM, north, and east coordinates, but the values were then converted to latitude and longitude, using Datum WGS84, to match the data from the GPS onboard the robot.

The transform system (TF) maintains the relationship between coordinate frames in a tree structure buffered in time. In other words, transformation configuration sets up the relationships between coordinate frames. Figure 6.11 shows the VRI4WD robot base with the sensors' relationship in a static transform publisher frames.

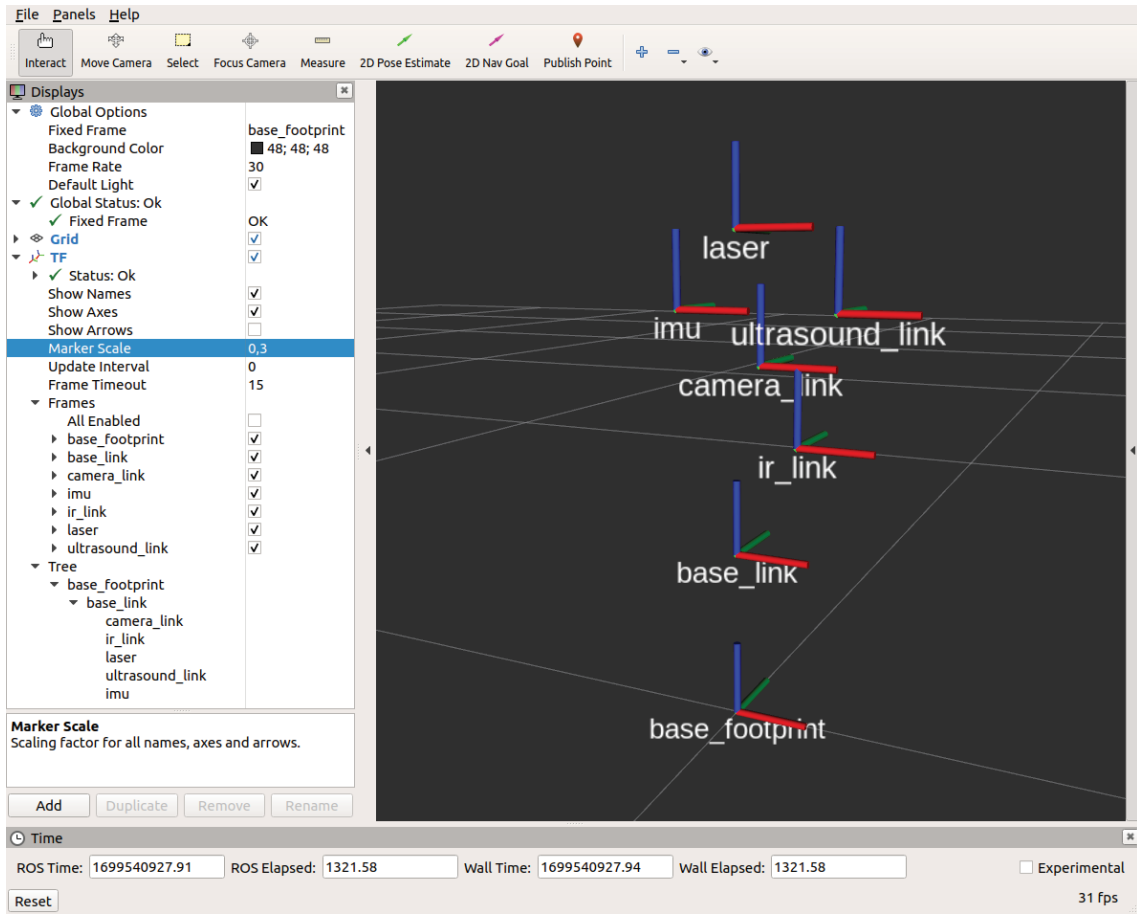


Figure 6.11: The RViz 3D visualization tool shows the setup of the relationships between coordinate frames in a tree structure. The TF VRI4WD platform indicated by the *base_footprint* and *base_link*, and the sensors relationship by the frames *ir_link* (3x infrared Sharp), *camera_link* (Kinect), *ultrasound_link* (5x URM37), *imu* (BNO055), and *laser* (laser scanner).

We recorded the data of the sensor map using the ROSBAG format. The ROSBAG is a set of tools for recording from and playing back to ROS topics. It is intended to be high-performance and avoids deserialization and reserialization of the messages (OpenRobotics, 2024). This way, it guarantees that all sensors can read topics simultaneously. The building of the database was segmented according to the trajectory points in Figure 6.12, following the clockwise direction starting at point A and ending at point F, indoor environment (AB and BF - *yellow path*) and outdoor environment (BC, CD, DE, and EB - *red path*), thus reducing the file size and also allowing the database to be used according to characteristics of interest of the environment.

The trajectories of the VRI4WD UFPR-MAP dataset were segmented between the following points, as represented by the type of environment, file, and date of data acquisition.

- Indoor A to B | **pathAB_vri4wd_ufpr-map_2025-03-20-18-07-27.bag** - Mar. 20, 2025
- Outdoor B to C | **pathBC_vri4wd_ufpr-map_2025-03-07-16-14-33.bag** - Mar. 07, 2025
- Outdoor C to D | **pathCD_vri4wd_ufpr-map_2025-03-07-16-20-18.bag** - Mar. 07, 2025
- Outdoor D to E | **pathDE_vri4wd_ufpr-map_2025-03-07-16-31-38.bag** - Mar. 07, 2025

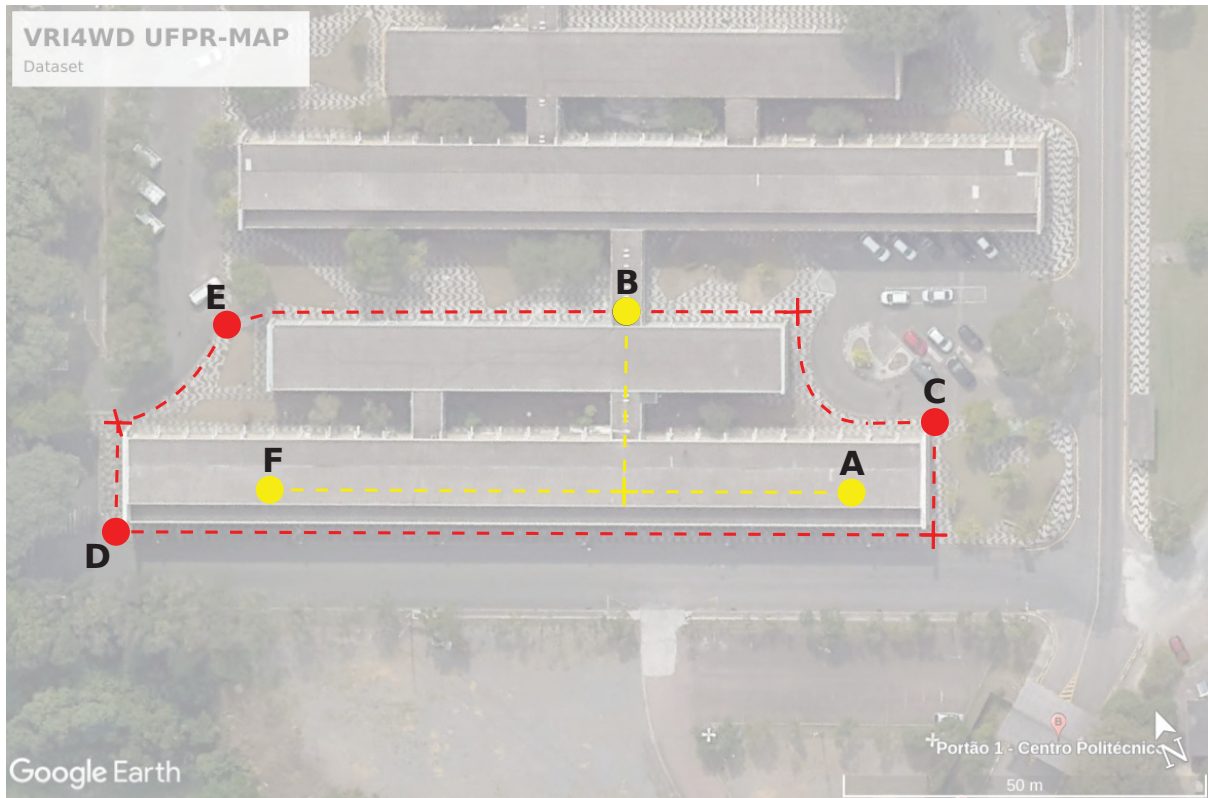


Figura 6.12: VRI4WD UFPR-MAP, paths taken on the map, Google Earth. Represents alphabetically the point-to-point trajectory of the robot in the map, indoor environment (AB and BF - *yellow path*), and outdoor environment (BC, CD, DE, and EB - *red path*).

- Outdoor E to B | **pathEB_vri4wd_ufpr-map_2025-03-07-17-18-43.bag** - Mar. 07, 2025
- Indoor B to F | **pathBF_vri4wd_ufpr-map_2025-03-20-17-57-04.bag** - Mar. 20, 2025

During the robot's movement in the indoor environment, we collected data at the same landmark points as ground truth. The architecture of the indoor floor allowed for standardizing the readings. Considering the irregularities and reference patterns of the outdoor environment, each position displaced the robot's position, depending on the orientation of the ground-truth (topographic points). File **robotPositionTTopographicPoints.csv** with the robot position table available at <https://github.com/VRI-UFPR/ufpr-map/tree/main/Ground-truth>.

6.2 CHAPTER CONSIDERATIONS

The work contributes significantly by building a sensor map dataset using ground-truth coordinates of landmarks obtained by precise geodetic and topographic methods. This UFPR-MAP dataset uniquely provides a meticulously georeferenced ground-truth for complex hybrid environments. Integrating data from a wide range of low-cost sensors and employing these advanced mapping methods offers an unprecedented resource for validating advanced localization and mapping algorithms, particularly those focused on sensor blend sets. The detailed documentation of its hardware, software (ROS topics), and the challenges overcome during its development further enhances its utility for the wider robotics community. This platform and dataset are expected to facilitate future research in robust autonomous navigation, especially in

challenging, real-world scenarios where diverse sensor inputs and high-precision ground-truth are relevant.

The experiments' direction and results used the VRI4WD platform or UFPR-MAP dataset, which was built to support this work and is available for future work.

The VRI4WD Mobile Robot is an *open-source project*, available at

<https://github.com/VRI-UFPR/vri4wd>

The VRI4WD UFPR-MAP *dataset* is available in

<https://github.com/VRI-UFPR/ufpr-map>

7 EXPERIMENTS

This chapter presents the experiments with the proposed cognitive blended sensors method, structured to allow segmented analysis across different levels of the perception-action cycle. The UFPR-MAP dataset provides a hybrid and realistic environment for evaluating multi-sensor perception. We implemented dedicated ROS packages to organize and format the information in the dataset according to the principles of the perception-action cycle, reflecting the cognitive architecture underlying the proposed model. To evaluate the approach's effectiveness, we conducted experiments utilizing sensors commonly employed for the localization of autonomous mobile robots, allowing for a comprehensive evaluation of the model in different environments.

7.1 THE DIRECTION OF EXPERIMENTS

The experiments guided the validation of the proposals raised by the hypotheses. In early experiments, Section 5.1, we tested and evaluated the localization and mapping of mobile robots using Gmapping SLAM in the Gazebo environment, using odometry data and a laser sensor. The influence of other approaches (Magrin and Todt (2016); Magrin et al. (2019); Magrin and Todt (2019b)) in developing a dedicated mobile robotics platform was considered, the concept of the *digital twin* used in Industry 4.0, regarding the proposal to use multi-sensors and the relevance of creating a digital model to train the learning algorithm more flexibly. The attention selection of a dataset evaluating indoor and outdoor environments to assist localization and mapping validates the proposal to use a hybrid environment, which is generally rare in the literature, as suggested by Magrin and Todt (2019a).

With the finalization of the dataset, UFPR-MAP, using the VRI4WD multi-sensor mobile robot, experiments were conducted to explore the hierarchical architecture at the three levels of sensors and motor of the learning perception-action cycle of cognitive blended sensors (Figure 4.3). The information used from the sensors in the development of the UFPR-MAP dataset is raw data; that is, the encoders only provide pulse counts and not odometry, sonars do not limit the value of the minimum and maximum range, Kinect presents direct information of the RGB image and depth - it is essential to run an algorithm to reduce information processing using, for example, edge detection algorithm, and the RF values show the attenuation of the signals and not their positioning. The objective of delivering raw data is to make the information available as received by the sensors without manipulating the data. As the UFPR-MAP dataset was made available in a ROSBAG package, it is only necessary to use the *play* command to read the topics.

7.1.1 CBS Perception-Action Cycle

Experiments were developed using the UFPR-MAP dataset to use the perception-action cycle. Packages were created for each level to format the information in the dataset using the concepts of the perception-action cycle of cognitive blend sensors. The packages used for experiments are available at: <https://github.com/VRI-UFPR/vri4wd>

7.1.1.1 *proprioceptiveSensors-locomotion (level 0)*

Some packages were developed to test the sensors in the dataset. In level 0, firstly, a package takes information in the topic */cmd_vel* for robot locomotion (*cbs_level0_locomotion.cpp*). In this case, the robot *subscribe* the topic */cmd_vel* and converts the linear and angular speeds into

PWM (Pulse-Width Modulation). Figure 7.1 shows data from the left and right encoders, accelerometer (x, y, and z), and gyroscope (x, y, and z), using the *cbs_level0_proprioceptivesensors.py* package.

```
[INFO] [1707314379.609864]: ***VRI4WD - PROPRIOCEPTIVE SENSORS*****
[INFO] [1707314379.614014]: Encoder (tick count) | L -3569 R -4034

[INFO] [1707314379.617285]: Accelerometer      | X -2.290000 Y -2.990000 Z -0.030000
[INFO] [1707314379.619846]: Gyroscope         | X -0.138889 Y 0.870000 Z 0.432222
[INFO] [1707314379.622203]: *****READING*****
```

Figura 7.1: VRI4WD proprioceptive sensors. The figure shows sensors reading data from encoders Left and Right in tick count and proprioceptive data from IMU: accelerometer (x, y, and z) and gyroscope (x, y, and z).

The encoder is given information by the */left_encoder* and */right_encoder* topics, only the pulse count, and the *cbs_level0_odometry.cpp* package publisher of simple Odom message where orientation.z is an Euler angle and publisher of full Odom message where orientation is quaternion (Figure 7.2). Figure 7.3 shows in RViz the odometry trajectory using encoder data from the **pathAB_vri4wd_ufpr-map_2025-03-20-18-07-27.bag** dataset. Due to a systematic error in pulse counting, it was necessary to a locomotion calibration on the right wheels by applying a multiplication constant of 1.014.

```
---
header:
  seq: 8604
  stamp:
    secs: 1707331961
    nsecs: 246911266
  frame_id: "odom"
child_frame_id: "base_footprint"
pose:
  pose:
    position:
      x: 7.50176853926
      y: 26.2896658129
      z: 0.0
    orientation:
      x: 0.0
      y: 0.0
      z: 0.975559711148
      w: 0.219734498849
  covariance: [0.01, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.01, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.01,
0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.1, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.1, 0.0, 0.0, 0.0, 0.0,
0.0, 0.0, 0.1]
twist:
  twist:
    linear:
      x: 0.372205450439
      y: 0.0
      z: 0.0
    angular:
      x: 0.0
      y: 0.0
      z: -0.904280696833
  covariance: [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0,
0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0,
0.0, 0.0, 0.0]
---
```

Figura 7.2: Odometry message. The figure shows the topic */odom_data_quat* where orientation is quaternion.

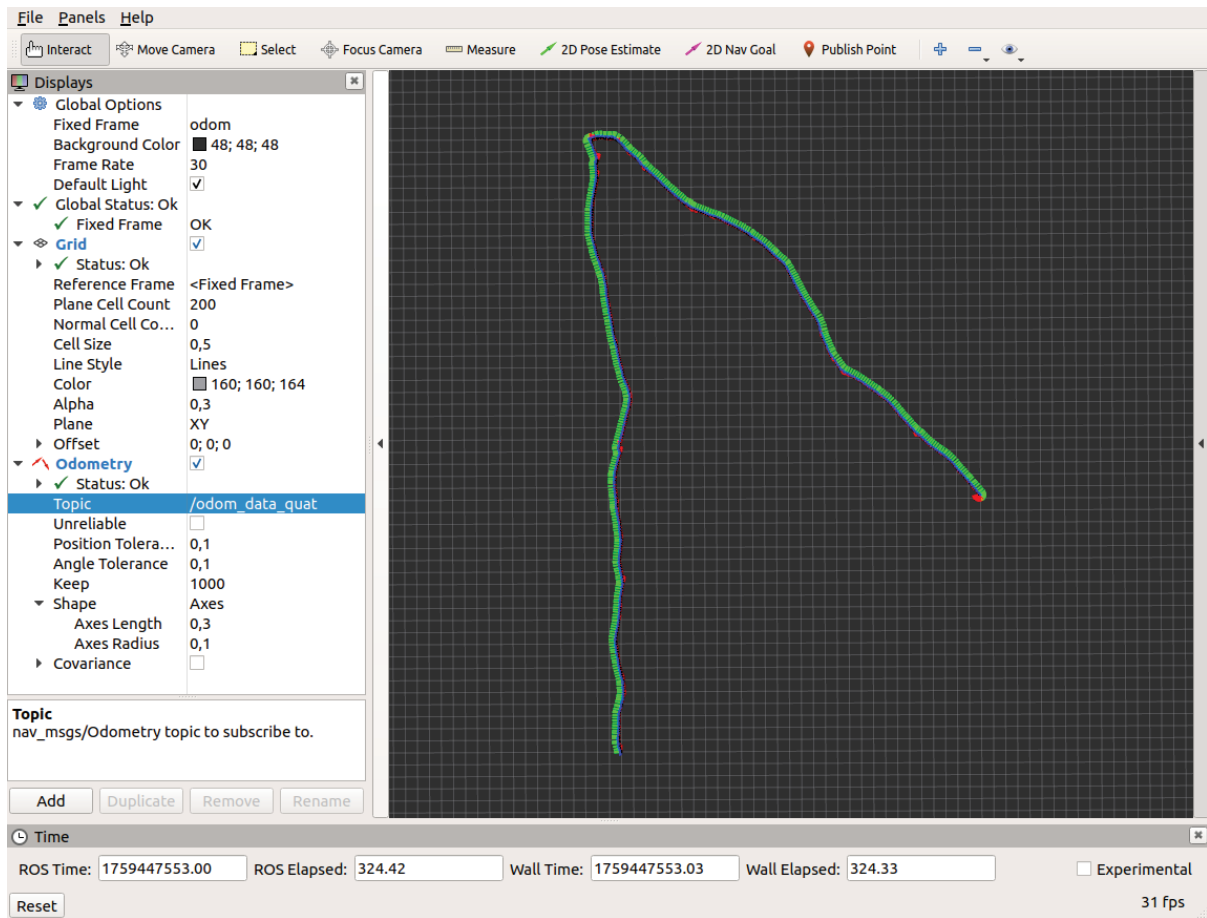


Figura 7.3: The figure shows in RViz the VRI4WD trajectory axes obtained from the raw encoder data using the odometry package, considering the `/odom_data_quat` topic.

The next hierarchical level addresses the exteroceptive sensors and motor control level, receiving information from the lower level.

7.1.1.2 *exteroceptiveSensors-motorControl (level 1)*

Hierarchical level 1 extracts features of the robot in the environment, available in the `cbs_level1_exteroceptivesensors.py` package, as shown in Figure 7.4, the distance sensors are competitive as they measure in the same direction left, front and right with IR and ultrasound modules. Positioning with the GPS compared to the Landmark (ground-truth) in latitude and longitude. The RF sensor, access point signal attenuation (RSSI), shows the raw values without normalizing the data with minimum and maximum values. The orientation is presented in degrees, using the *Yaw* of Euler angles, values read from the topic `/imu/data`. A proportional constant K_p was added to assist in motor control, using as a reference the information from the encoders, proprioceptive sensors, in the topics `/left_encoder` and `/right_encoder` using the package `cbs_level0_locomotion.cpp`.

For the Xbox 360 Kinect camera `/camera/rgb/image_mono` (Vision), the topic `/camera/rgb/image_mono` was used. To reduce the amount of image information, the Canny edge detector algorithm from OpenCV - Canny Edge Detection was applied through the topic `cbs_level1_cannykinect.py`. Figure 7.5 shows an example of the monochromatic image on the left and, in the frame on the right, the output with the Canny algorithm.

```

[INFO] [1759443400.490500]: ***VRI4WD - EXTEROCEPTIVE SENSORS*****
[INFO] [1759443400.494779]: IR (cm) | L 2843.67 F 9521.66 R 141.15
[INFO] [1759443400.498753]: Ultrasound (cm) | L 1026.00 F 1022.00 R 122.00

[INFO] [1759443400.502054]: Point | P073
[INFO] [1759443400.505786]: Landmark | Latitude -25.452877 Longitude -49.233193
[INFO] [1759443400.509252]: GNSS | Latitude -25.453596 Longitude -49.232792

[INFO] [1759443400.512431]: Orientation | Yaw Degree 359.895217

[INFO] [1759443400.515318]: RF Sensor | W1 31 (VRI4WD_WLAN * Embedded * Signal (-dBm))
[INFO] [1759443400.518247]: RF Sensor | W2 75 W3 84 W4 88 W5 92 W6 92
[INFO] [1759443400.521482]: RF Sensor | W7 87 W8 86 W9 87 W10 89 W11 87
[INFO] [1759443400.524645]: RF Sensor | W12 89 W13 89 W14 89 W15 0 W16 0
[INFO] [1759443400.528200]: RF Sensor | W17 0 W18 0 W19 0 W20 0
[INFO] [1759443400.531713]: *****READING*****

```

Figura 7.4: VRI4WD exteroceptive sensors. The figure shows sensors reading data from IR and ultrasound (Left, Front, and Right) in centimeters, latitude and longitude from GNSS, orientation in degree, and RF sensors from access points available in the UFPR campus. The dataset presents the latitude and longitude from and ground-truth (landmark) and point in the map for reference.

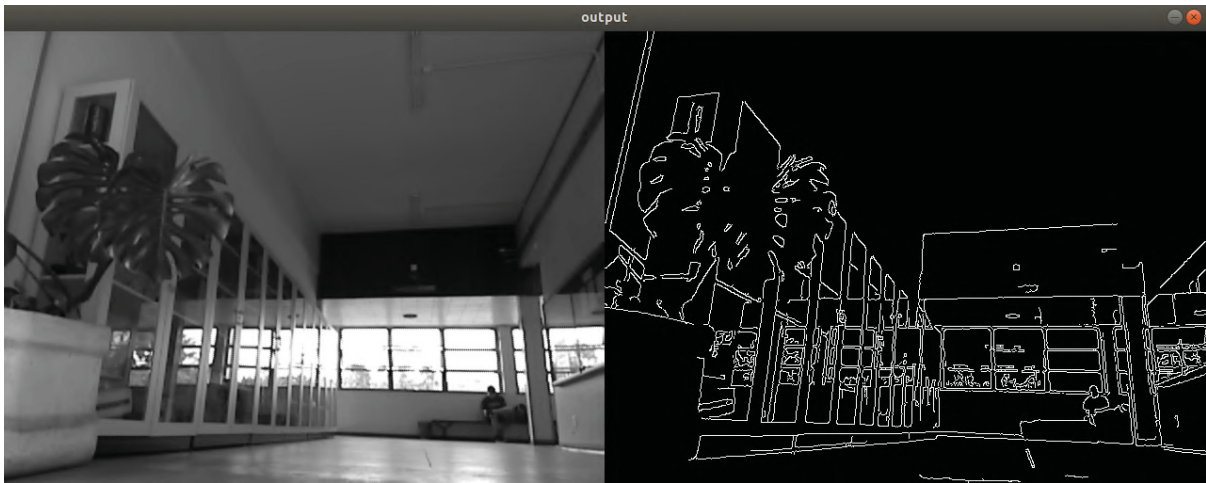


Figura 7.5: Camera using Kinect. The image on the left shows the topic `/camera/rgb/image_mono`, and on the right, the output with image processing using the OpenCV Canny Edge Detection algorithm.

Figure 7.6 shows the output of the `/scan` thread on RViz. Reference image to the *AB* path of the dataset, that is, laser mapping in an indoor environment.

7.1.1.3 *sensorFusion-pathPlanning (Level 2)*

The last level of the architecture allows the integration of other levels *proprioceptiveSensors* (level 0) and *exteroceptiveSensors* (level 1), that is, sensor fusion. Considering SBS, sensor fusion can occur in the same set or even with sensors from different sets. Experiments were carried out to validate the fusion of some sensors in a complementary and collaborative way.

Due to the problems described in the validation of the inertial measurement unit (IMU), the BNO055 unit was selected, which intrinsically runs a fusion of sensors combining an accelerometer, gyroscope, and magnetometer to calculate values that describe the absolute orientation of the device in the world, in addition to delivering the values in quaternion or Euler angles (Bosch Sensortec, 2021). Figure 7.7 shows the axes reference to Euler angles (roll (red axis), pitch (blue axis), and yaw (green axis) reference the collaborative fusion of accelerometer and gyroscope sensors.

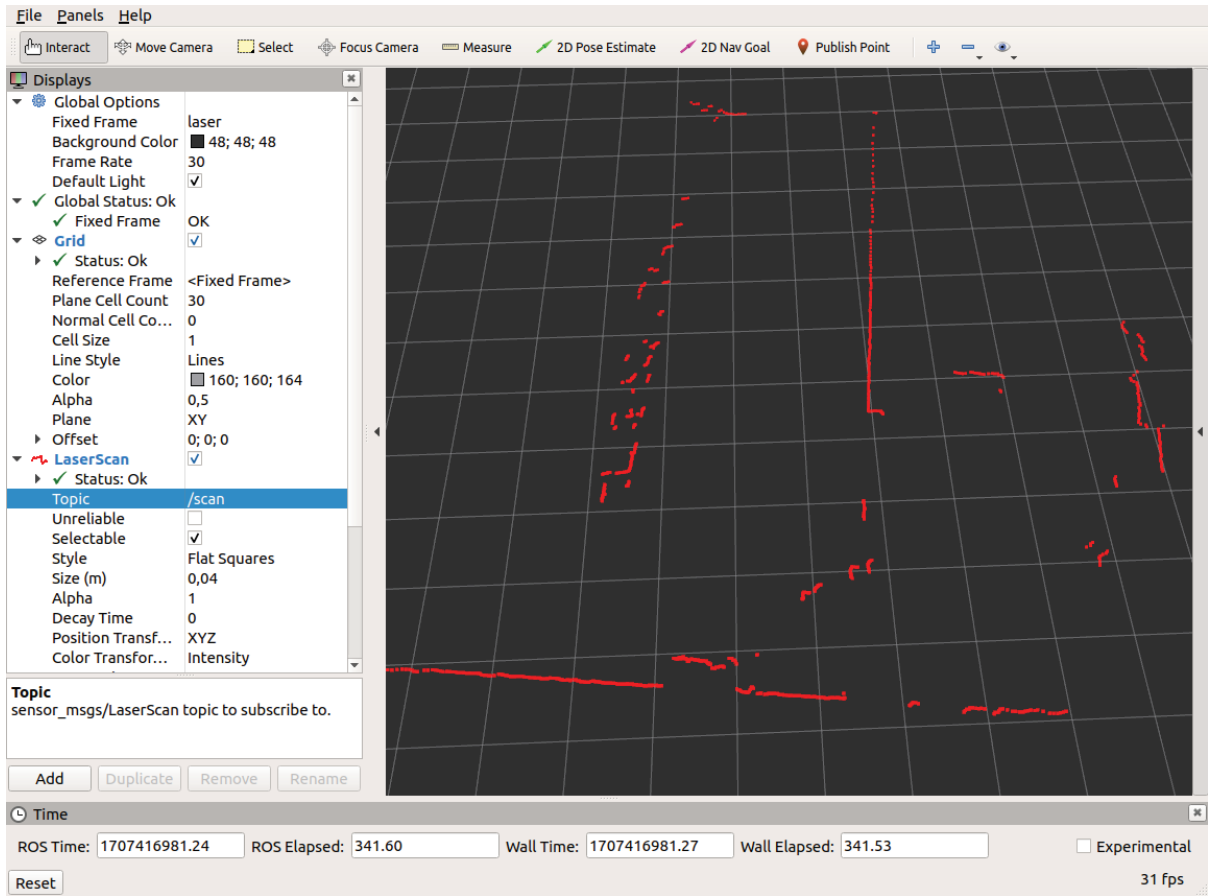


Figura 7.6: Laser scan mapping. The image shows in RViz screen laser scan topic `/scan` from *AB* path of the dataset. VRI4WD Robot using laser scanner Slamtec RPLIDAR-A1 with 12m range in 360°.

The VRI4WD platform provides information from 12 different types of sensors, divided into five blends with different characteristics to complement the cognitive process. The proposed approach is validated by comparing its performance on the dataset with established methods and their primary sensors. These comparative tests include: outdoor positioning using GNSS; indoor positioning using the kNN fingerprint technique with RF sensors; place classification based on deep learning vision algorithms; and Gmapping, which utilizes particle filters for encoder and laser sensor fusion.

7.2 RESULTS OF EXPERIMENTS

To evaluate the proposed model, some experiments were analyzed using sensors generally used to assist the localization of autonomous mobile robots in indoor or outdoor environments.

7.2.1 Landmarks *versus* GNSS

Using *rqt_multiplot*, the positioning of the landmark points (ground-truth) and the GNSS positioning were compared. Figure 7.8 shows 2D plots of the outdoor paths B to C and C to D, and Figure 7.9 shows 2D plots of the outdoor paths D to E and E to B. Positioning of the GNSS module, even outdoors for the environment's architecture, close to walls, did not favor the location. The influence of the multipath effect on the outdoor E to B path was the most displaced by the presence of walls on both sides, so on the outdoor C to D path, the variation was smaller, as even less than 2 meters from the wall, there is a more open field for reception of the positioning.

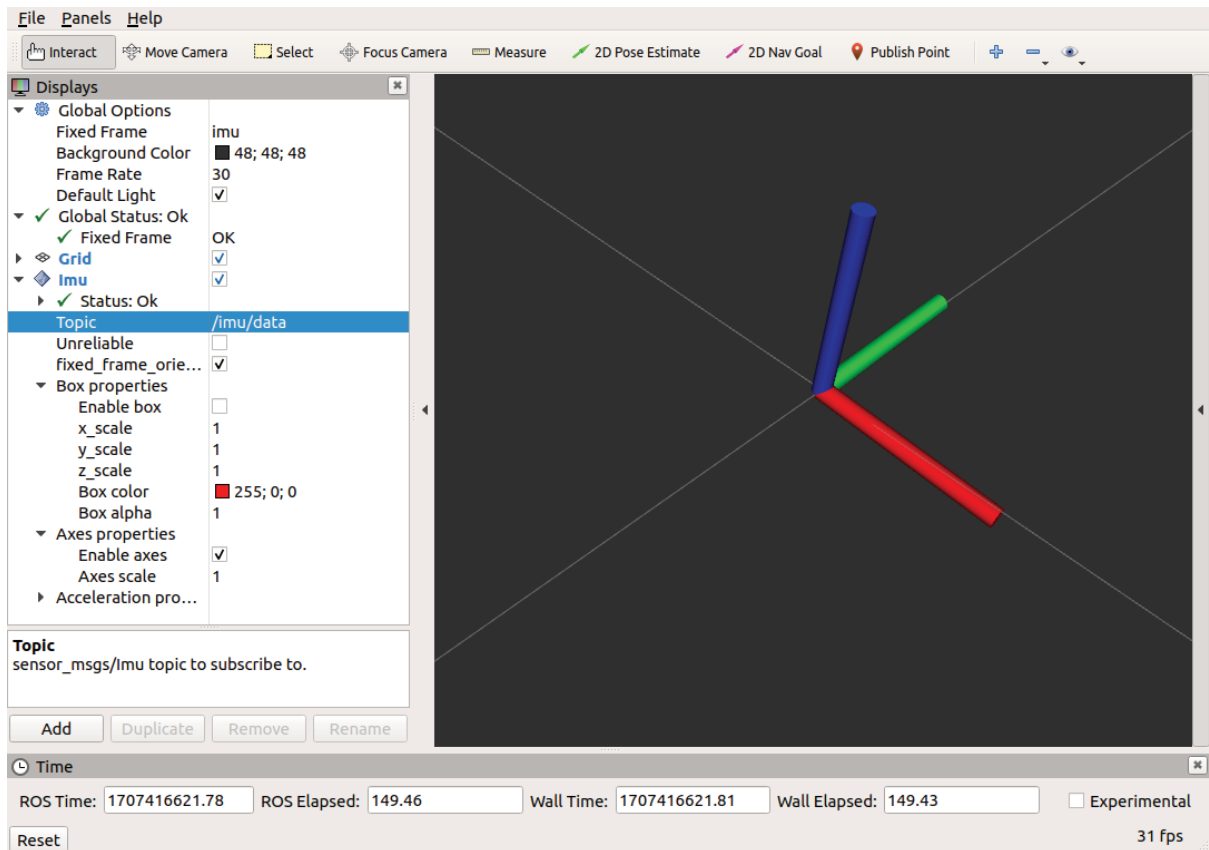


Figura 7.7: IMU frame orientation. The image shows the fixed frame orientation in RViz screen Imu topic `/imu/data`. The axes of reference to Euler angles are roll (red axis), pitch (blue axis), and yaw (green axis).

Figure 7.10 shows the inconsistency of the GNSS positioning for indoor environments, even for measurements that were close to windows and the GNSS module normally receiving positioning, the multipath effect was high for the environment in relation to the outdoors (Figure 7.8 and Figure 7.9).

To check the GNSS positioning paths in the environment, a walk was carried out along the paths of the UFPR-MAP dataset with a Garmin Forerunner 245 watch configured with GPS + GLONASS. Using the GPX (GPS Exchange) format file created by Garmin watch, the path with the ground-truth points was plotted on Google Earth map. Figure 7.11 shows the outdoor trajectory and points on the Google map in red. The location allows an approximate positioning of the surroundings of the geology building but with significant positioning errors at the reference points. Considering the difficulty of indoor positioning of the GNSS system, Figure 7.12 shows the indoor trajectory with a large displacement of the points on the Google map in yellow.

Concluding the evaluation of landmark vs. GNSS, the use of GNSS modules to obtain the positioning of autonomous mobile robots guarantees the geographic coordinates in latitude and longitude of the environment, but with considerable error due to the architecture of the building. Robot positioning using GNSS does not individually address the robot pose, but the data can be integrated into a sensor fusion.

7.2.2 Environment Classification

The GNSS positioning results in Section 7.2.1 confirmed that GNSS alone is insufficient to localize the robot, even in outdoor environments, due to the robot's proximity to the building. As the context of the work is cognitive sensor fusion in hybrid environments, environment

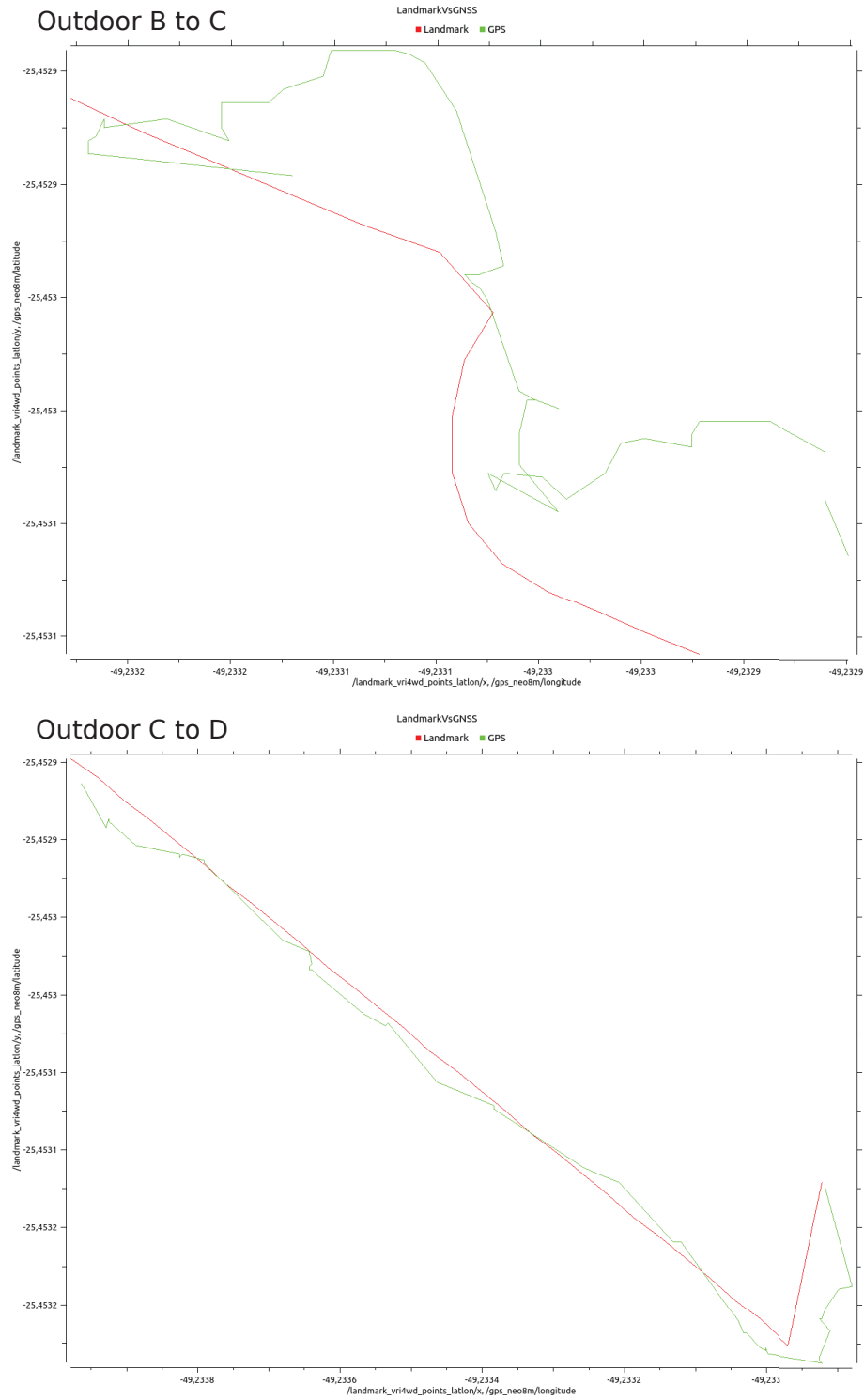


Figura 7.8: 2D plots of the outdoor path D to E and outdoor path E to B. The image shows the ground-truth (landmark) in red and the robot positioning (GNSS) in green, considering the longitude on the x-axis and the latitude on the y-axis.

classification was validated using image segmentation with Kinect vision data. Mobile robot localization using the WLAN structure with RF sensor data was evaluated for path classification and environment segmentation to complement the environment classification.

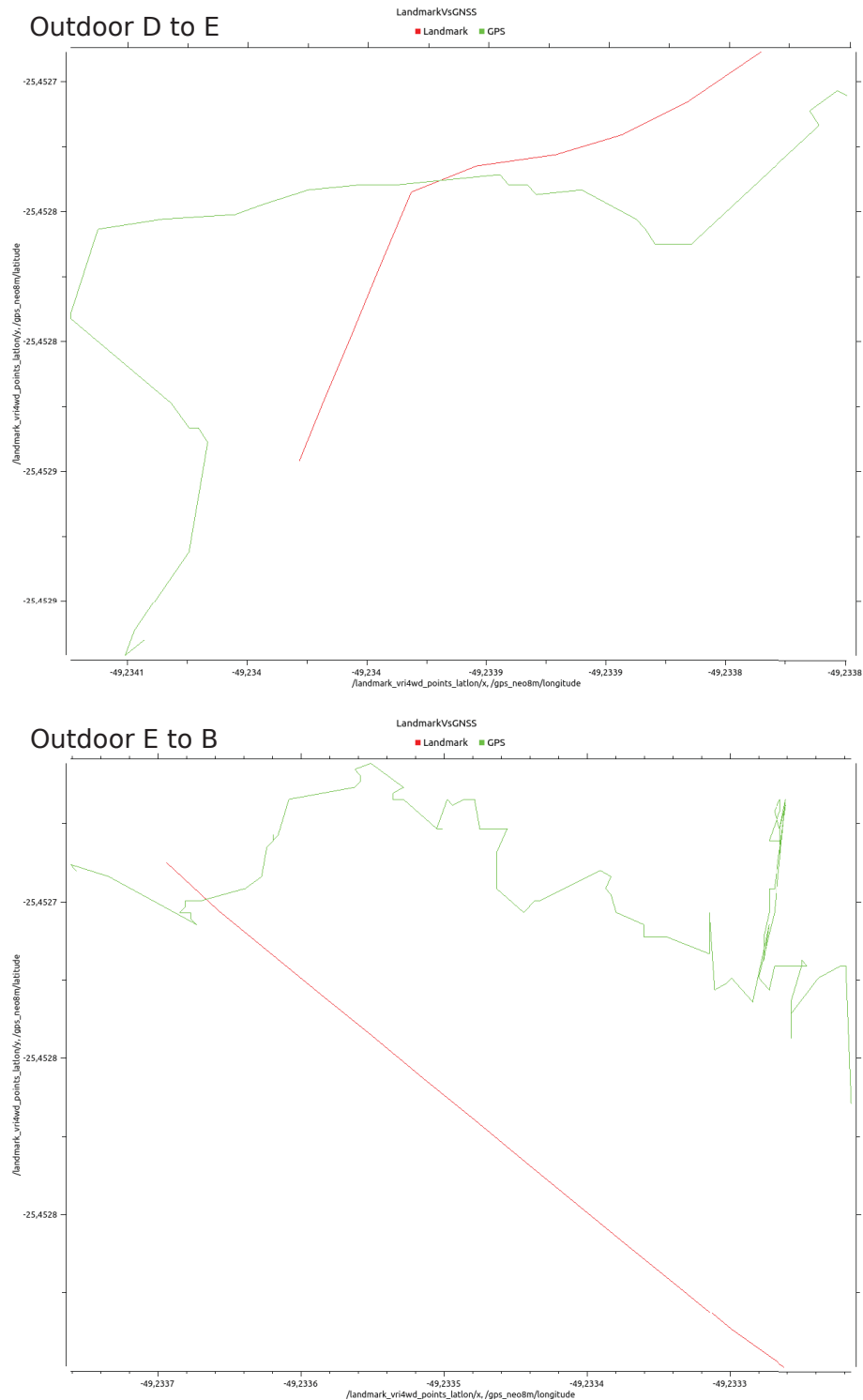


Figura 7.9: 2D plots of the outdoor path B to C and outdoor path C to D. The image shows the ground-truth (landmark) in red and the robot positioning (GNSS) in green, considering the longitude on the x-axis and the latitude on the y-axis.

7.2.2.1 Image segmentation

Starting from the experiments and classification of places with the NCLT Dataset, a deep learning algorithm was used to classify environments with the Keras *framework*. Image segmentation was performed with the extract images from a ROSBag *bag_to_images.py* package through the ROS topic */camera/rgb/image_color*. Figure 7.13 shows a sample image of each path

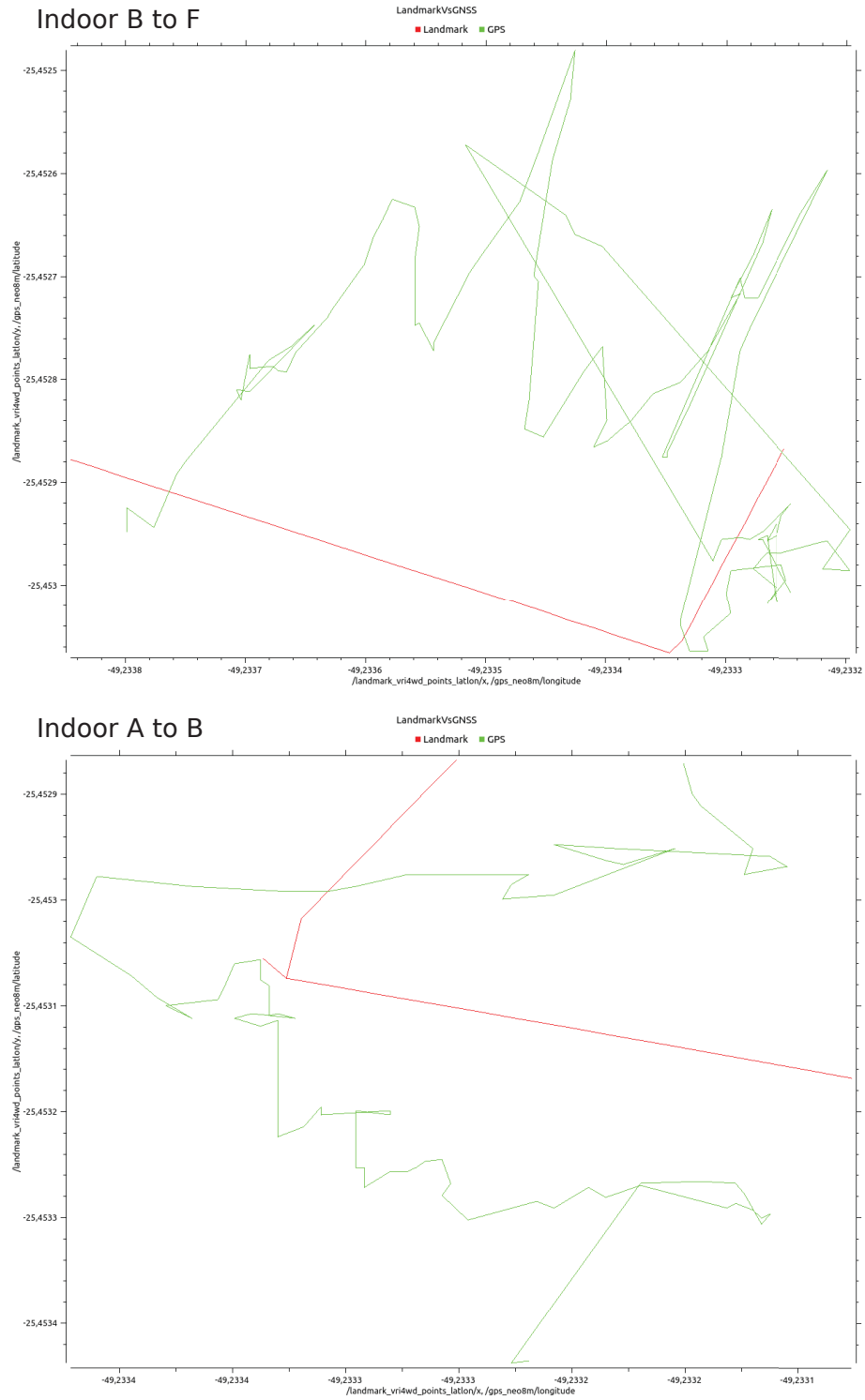


Figure 7.10: 2D plots of the indoor path B to F and outdoor path A to B. The image shows the ground-truth (landmark) in red and the robot positioning (GNSS) in green, considering the longitude on the x-axis and the latitude on the y-axis.

from the UFPR-MAP dataset used to validate the classification of indoor or outdoor environments, according to Figure 6.12.

Model training used 1,527 files belonging to 2 classes (indoor and outdoor) and 305 files for validation. To arrive at a suitable result for the binary classification, more epochs were necessary to leave 50%

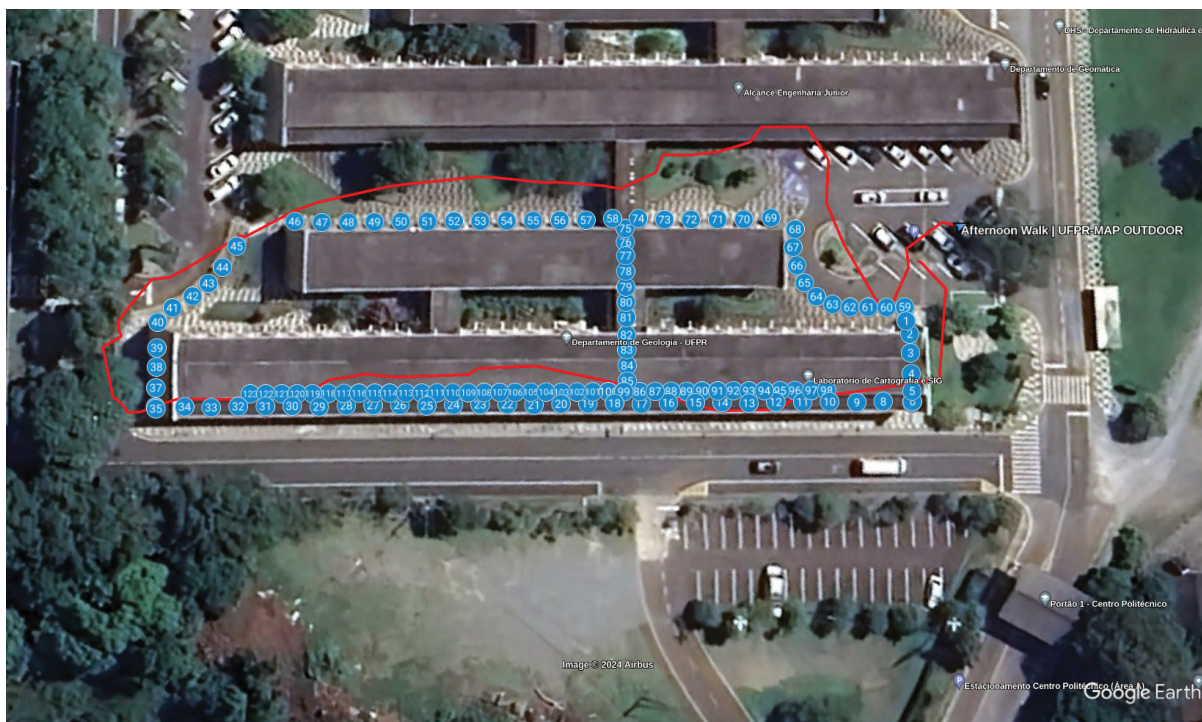


Figura 7.11: Afternoon walk in the outdoor paths using Garmin Forerunner 245 watch. Google Earth map shows the landmark points in blue, and the GPX format file plots the outdoor trajectory in red.

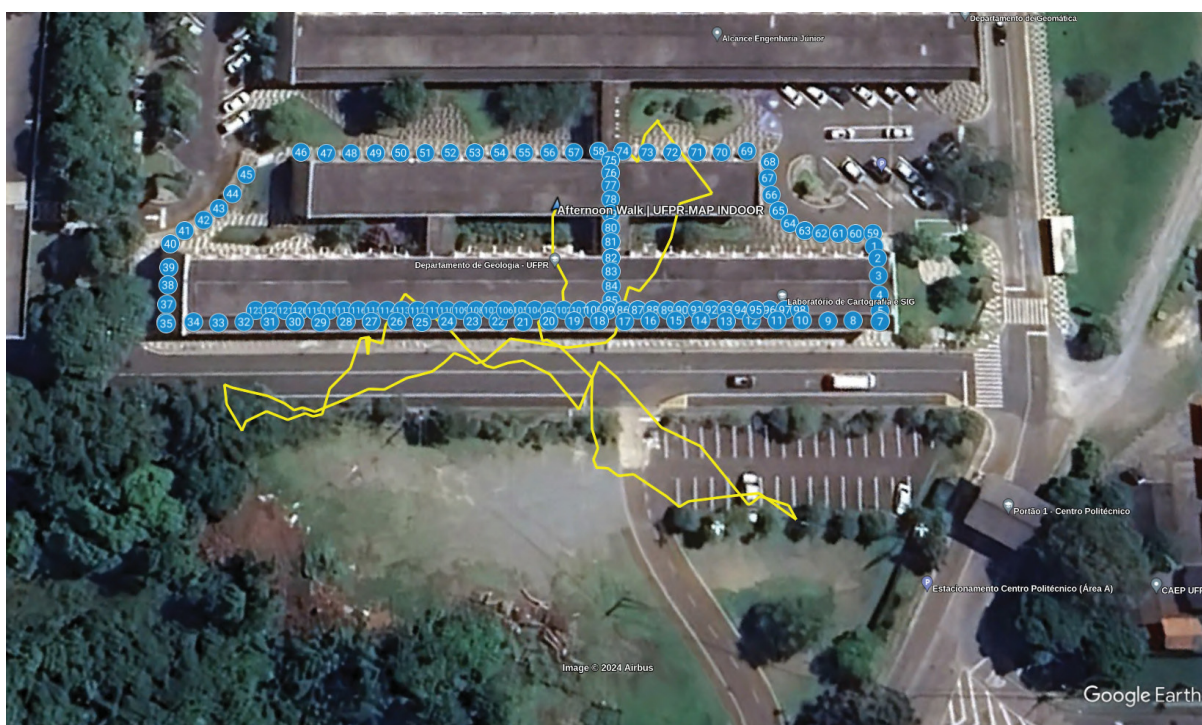


Figura 7.12: Afternoon walk in the indoor paths using Garmin Forerunner 245 watch. Google Earth map shows the landmark points in blue, and the GPX format file plots the indoor trajectory in yellow.

To validate the environment classification, an image from each path was selected (Figure 7.13). The objective of the classification would be cognitive fusion with the sensor blend sets, choosing an indoor or outdoor environment for a better selection of multi-sensors. Table 7.1 shows the environment recognition rate in each path in the VRI4WD UFPR-MAP



Figura 7.13: Image classification indoor or outdoor environment. The image shows the following paths in each environment; they are (a) indoor path-AB, (b) indoor path-BF, (c) outdoor path-BC, (d) outdoor path-CD, (e) outdoor path-DE, (f) outdoor path-EB.

dataset. It is worth mentioning that the images extracted from the Kinect camera, topic `/camera/rgb/image_color`, are raw and low-resolution data 640 x 480 pixels.

Tabela 7.1: Validation of new data. The table shows the 100% recognition rate for all outdoor samples with a minimum error for indoor environments.

Image Color	Indoor	Outdoor
ufpr-map_20230830s001ab_image_color0170.png	99.98%	0.02%
ufpr-map_20230830s001bc_image_color0004.png	0%	100%
ufpr-map_20230830s001cd_image_color0079.png	0%	100%
ufpr-map_20230901s001de_image_color0030.png	0%	100%
ufpr-map_20230901s001eb_image_color0132.png	0%	100%
ufpr-map_20230901s001bf_image_color0280.png	99.94%	0.06%

The images used to classify environments are available at: https://github.com/VRI-UFPR/ufpr-map/tree/main/images_color

7.2.2.2 WLAN localization

Using the UFPR-MAP dataset, data was collected from the RF sensor of the VRI4WD robot, which takes advantage of the WLAN structure of the campus and laboratories along the trajectory to read the RSSI of the APs, a total of 19 APs. The environment classification uses segmentation of path, environment, and landmark. The data was extracted from the bag file using ROS Topics `$rostopic echo -b dataset_vri4wd_ufpr-map_20230830s001ab.bag -p /rssi1_wlan > data.csv`. A key aspect of our methodology was using the kNN (K-Nearest Neighbors) supervised learning algorithm. This algorithm was instrumental in creating a ROS package for path and environment localization, `cbs_level2_wlanlocalization.cpp`, which publishes the topics `/wlan_path` and `/wlan_environment`. The paths correspond to Figure 6.12, considering AB, BC, CD, DE, EB, and BF. The environment indication only segments paths when changing characteristics, entering another corridor or side of the building.

According to Wettscheck and Aha (1995), many algorithms derived from the kNN classifier are sensitive to the relevance of each feature, and the method used to reduce the sensitivity of the kNN similarity function is feature weighting. The weighting for each feature represents values between 0 and 1, so 0 (zero) is considered an unavailable value, and 1 (one) is an available value in the comparison calculation. The equation 7.1 represents the weighted Euclidean distance calculation, where *train* is the position in the training vector (`wlan_rssiMap_pathEnvironment_rawData.csv`), *sensor* the positioning vector and ω the weighted value for each RSSI information.

$$d = \sqrt{\sum_{i=1}^n (train_i - sensor_i)^2 * \omega_i} \quad (7.1)$$

For validation, the weighted Euclidean distance in k-NN was used to exclude or reduce the relevance of some data without different characteristics between the vectors. The features RSSI1 (robot access point) and the features between RSSI12 and RSSI19 were represented as 0 *zero* weight as they did not present similarity for the classification, making the data unavailable for environment localization. Weight 1 *one* was used for other characteristics.

Table 7.2 provides a clear view of the recognition rate of the path and environment. MAP indicates validation with all trajectories, and PATH validation with only the segment considering a hierarchical location. The algorithm first selects the path and then locates the environment. We encountered significant challenges when using the dataset for localization using all 123 UFPR-MAP landmarks, as the recognition rates for the CD, DE, and EB routes were less than 50%. This underscores the complexity of our task and the need for further research in this area.

The segmented dataset used to classify paths, environments, and landmarks is available at: https://github.com/VRI-UFPR/ufpr-map/tree/main/wlan_rssimap

Concluding the evaluation of environment classification, the results presented in Table 7.1 show an environment recognition rate suitable for classifying indoor or outdoor environments. Classification of the environment using the image is extremely useful to assist the positioning sensor blend set, as the GNSS can be allocated to outdoor positioning, where it has the best response in location. WLAN localization results show a recognition rate of over 90% for indoor segmentation, aiding sensor fusion to localize the mobile robot pose at the sensor fusion level.

Tabela 7.2: The table shows the validation for WLAN localization using data from the UFPR-MAP dataset. Using the k-NN algorithm, setting k=3, the recognition rate for the path applying the entire MAP was 96%, but for the environment using MAP as the whole, it was less than 50%. The MAP segmentation in PATH was validated for the environment, presenting a less than 80% recognition rate and the remainder greater than 90%.

UFPR-MAP		Recognition Rate
Path	MAP	96%
	MAP	49%
Environments	PATH-AB	95%
	PATH-BC	100%
	PATH-CD	100%
	PATH-DE	80%
	PATH-EB	100%
	PATH-BF	94%

7.2.3 Simultaneous Localization and Mapping

In this experiment section, the sensor fusion evaluation was conducted using the SLAM (Simultaneous Localization and Mapping) algorithm GMapping as a validation platform. GMapping was employed to analyze the integration performance between proprioceptive and exteroceptive sensor data from the VRI4WD, which was developed within the objective of this work. Specifically, the process evaluated the consistency between the odometry estimate, calculated from the wheel encoders (tick count), using the odometry package (*cbs_level0_odometry.cpp*), and the environmental measurements provided by the laser scanner. The experiments were performed on the UFPR-MAP dataset, both indoor and outdoor paths. We used the Gmapping package (*/cbs_level2_gmapping_ufpr-map*) to understand the limitations of using different sensors. Thus, the quality and consistency of the map generated by GMapping are used as a quantitative and qualitative indicator of the sensor fusion, validating the performance of the odometry system in a real-world application scenario.

Figures 7.14 e 7.15 show the map results generated by Gmapping SLAM in indoor environments. In addition to the map, RViz was configured to display laserscan, odometry, and transformation (TF) topics to visualize the map construction and particle filter corrections in sensor fusion between the encoders and the laser scanner.

The experiments using low-cost laser scanners combined with odometry based on encoder pulses were not initially expected to yield a high-quality map representation in outdoor environments. This limitation is mainly due to two factors: the excessive ambient light, which interferes with the performance of laser sensors, and the irregular terrain, which affects the platform's mobility and leads to considerable wheel slippage. Additionally, the limited scanning range of the laser and the building's geometry caused the system to primarily detect the wall on the right-hand side during the mapping trajectory. Figure 7.16 shows that the right wall was accurately reconstructed, even revealing the structural details of the pillars used for the landmark (tag) point. Conversely, Figure 7.17 shows a failure of Gmapping to maintain localization consistency, resulting in an incomplete and distorted map reconstruction.

Regarding the results of SLAM in hybrid environments, sensor fusion plays a key role in improving performance. The combination of LiDAR and wheel encoders was validated in indoor environments with flat surfaces and outdoor areas with limited terrain irregularities and minimal interference from lighting conditions. In addition to SLAM, a preprocessing stage based on an Extended Kalman Filter (EKF) to fuse odometry and IMU data could further enhance

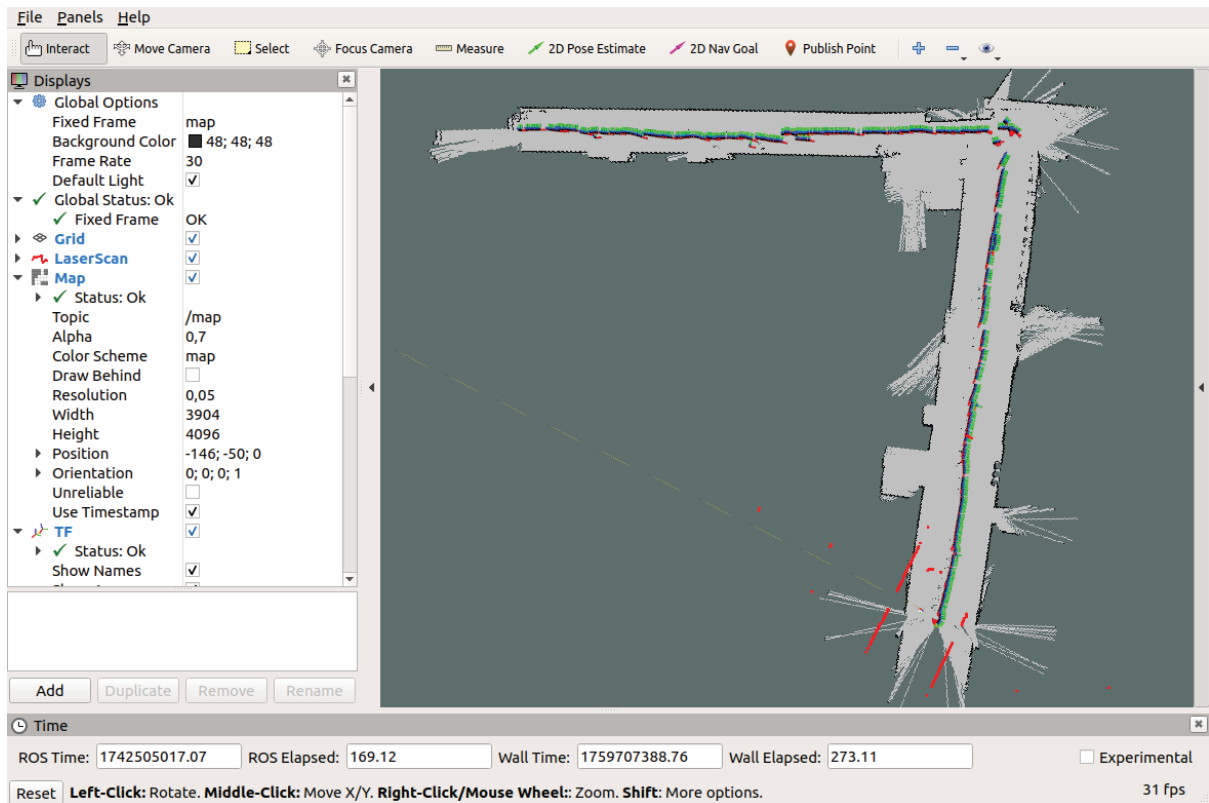


Figura 7.14: The image shows the 2-D occupancy grid map generated by the GMapping SLAM package, visualized in RViz. The map was built from odometry and laser scanner data collected on the PATH-AB of the UFPR-MAP dataset.

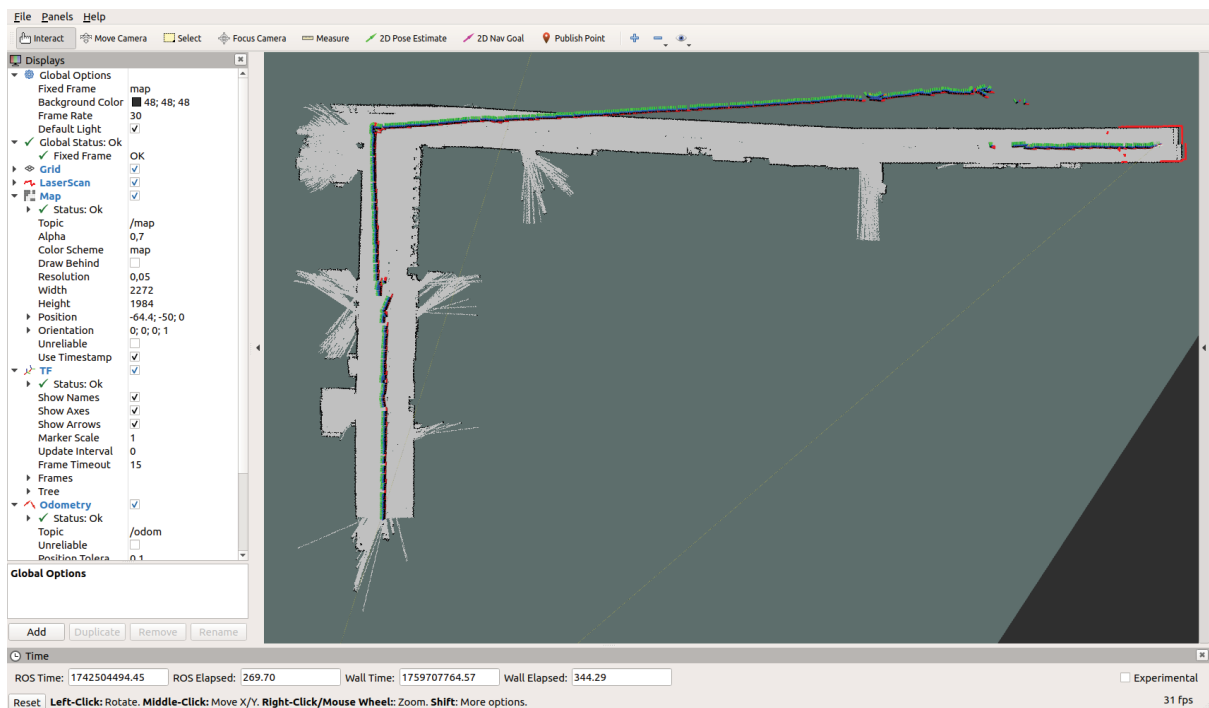


Figura 7.15: The image shows the 2-D occupancy grid map generated by the GMapping SLAM package, visualized in RViz. The map was built from odometry and laser scanner data collected on the PATH-BF of the UFPR-MAP dataset.

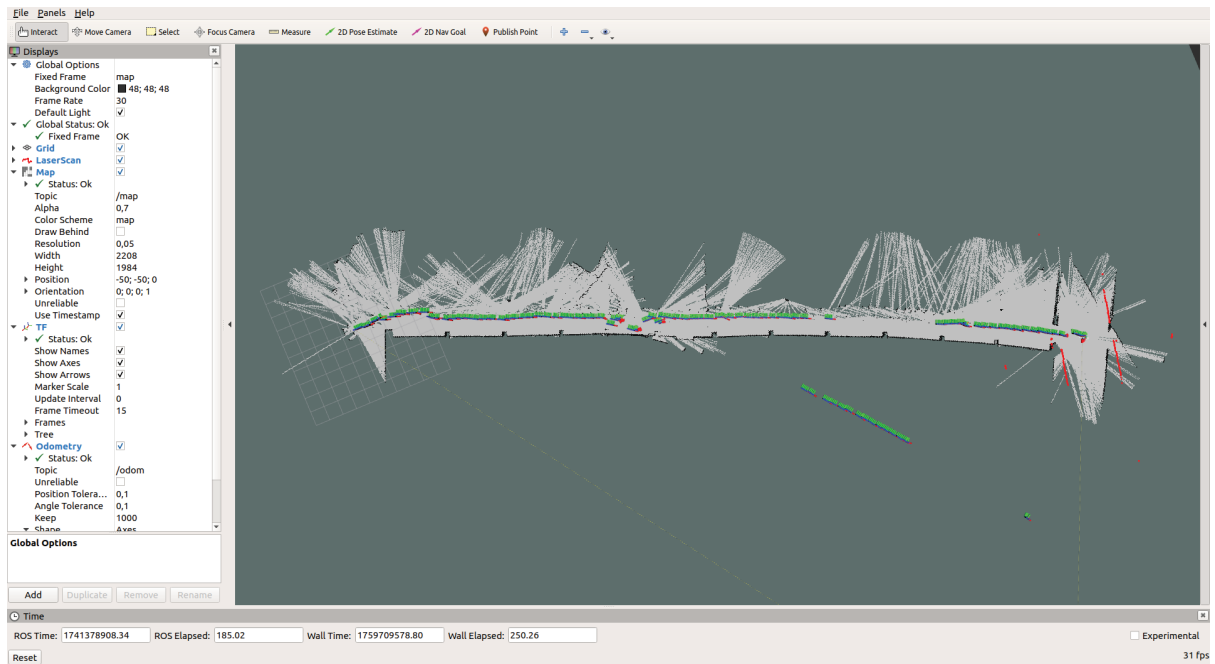


Figura 7.16: The image shows the 2-D occupancy grid map generated by the GMapping SLAM package, visualized in RViz. The map was built from odometry and laser scanner data collected on the PATH-EB of the UFPR-MAP dataset.

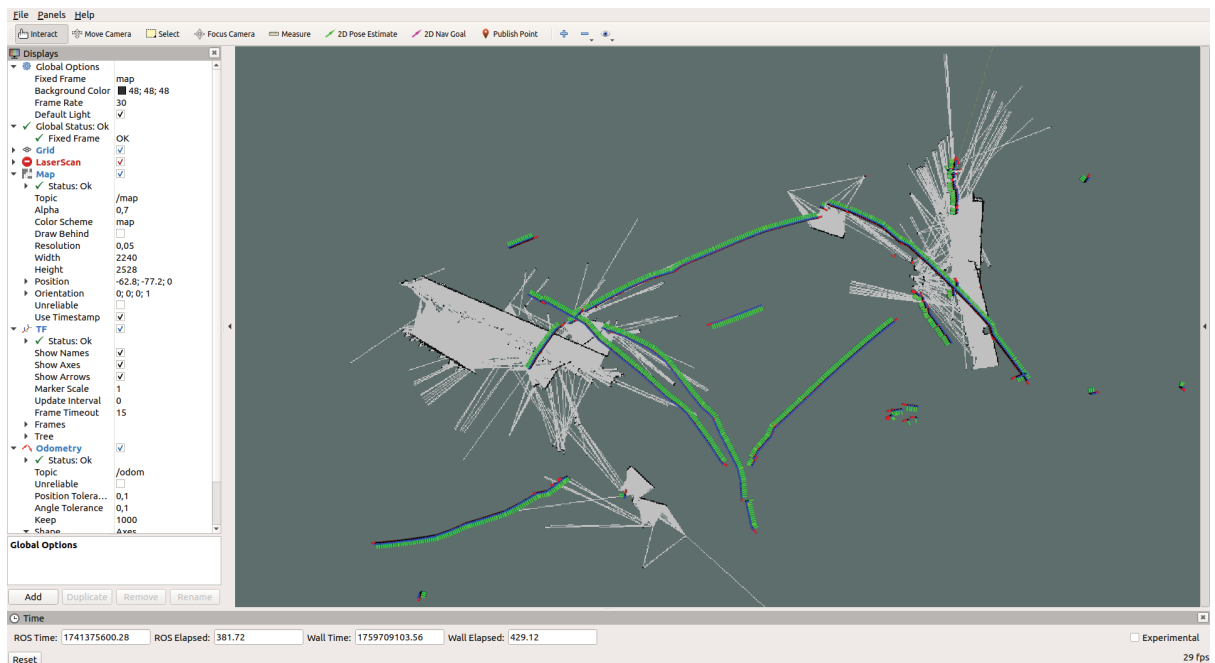


Figura 7.17: The image highlights a failure of GMapping to maintain localization consistency during the trajectory, resulting in an incomplete and distorted map reconstruction. The map, visualized in RViz, was built using the GMapping algorithm from odometry and laser scanner data collected on the PATH-CD of the UFPR-MAP dataset.

localization in trajectories subject to higher levels of irregularity. However, this was beyond the scope of this study, which focused on highlighting both the limitations and the strengths of applying a SLAM algorithm across distinct environmental conditions.

7.3 CHAPTER CONSIDERATIONS

The experiments presented in this chapter provided a segmented and structured evaluation of the proposed cognitive sensor fusion model, emphasizing its alignment with the perception-action cycle. We organized many types of raw sensor data from the UFPR-MAP dataset into hierarchical levels and processed them using custom ROS packages to reflect the model's cognitive architecture. The availability of a dataset using a real mobile robot with sensor sets for different sensor fusion tasks, along with geodetic ground truth, was also important, as was the VRI4WD UFPR-MAP dataset. This dataset also allowed for the exploration of mapping in real environments. Although some results exhibited measurement errors, these discrepancies were consistent with the known limitations of individual sensors, validating the importance of considering sensor constraints in real-world applications. This highlights the crucial role of cognitive sensor fusion in overcoming such limitations by integrating multiple complementary sensor modalities. Therefore, the experimental outcomes not only demonstrate the effectiveness of the proposed model but also highlight the necessity of multi-sensor approaches to achieve robust perception and navigation.

8 CONCLUSION AND FUTURE WORK

This thesis addressed the critical challenge of mobile robot localization in indoor and outdoor environments, where existing sensor fusion techniques often lack adaptability and comprehensive validation datasets. This work was aimed at assisting in the localization and mapping of autonomous mobile robots in a hybrid environment, as there is a gap in the published works related to the transition between indoor and outdoor environments during mapping. Furthermore, different types of sensors or even concepts are directed towards the selection of sensors and algorithms in each kind of environment. Identifying the tasks of the sensor sets or even the perception-action cycle of an autonomous robot is a great challenge for the fusion of multiple sensors.

To this end, the study of cognitive systems and the architecture of the mind led to the organization of an analogy of the perception-action cycle of the cognitive system and the development of a hierarchical sensor fusion method, Cognitive Blended Sensors (CBS). This thesis advances sensor fusion with the Sensor Blend Sets (SBS) model, presenting a set of multi-sensors for different tasks such as state, movement, distance, mapping, and positioning. Research on multi-sensor fusion and cognitive systems led to a systematic review, the application of a neural network method for hierarchical sensor fusion, and validation of the simulation platform and the digital twin as a proposed learning tool. The study of the main sensors used in wheeled mobile robots and their application in different public datasets, such as KITTI, Oxford RobotCar, and NCLT, highlighted the need to develop a dedicated robotic platform for sensor mapping in a hybrid environment. In this context, we created a dataset with sensor mapping of path transitions from indoor to outdoor environments, with point coordinates obtained using geodetic and topographic methods, as differential ground truth. It is not usual to find georeferenced datasets in the literature; this aspect constitutes an important feature of our proposal.

These contributions offer a novel theoretical framework for sensor management in complex scenarios and an open-source VRI4WD robot platform and the UFPR-MAP dataset, significantly lowering the barrier for future research in robust, low-cost autonomous systems. The fusion of multi-sensors deepens our understanding of the perception-action cycle used in mobile robotics for a model that utilizes the concept of cognitive systems with Cognitive Blend Sensors. The work significantly contributes to the development of the VRI4WD robot, which uses many types of sensors for indoor and outdoor environments, to meet the proposed Sensor Blend Sets of perception tasks.

While this work demonstrates significant advancements, current limitations include the level of cognitive learning, so the level of attention could move towards a dynamic selection of the Sensor Blend Sets according to the type of environment and challenges in the localization and mapping of the robot.

Building upon these findings, future work could focus on cognitive attention learning in the cognitive blended sensors method, adaptively selecting blended sensor sets. Developing a digital twin of the UFPR-MAP scenario and the VRI4WD robot could aid in developing new sensor fusion methods, primarily based on reinforcement learning. We can expand the UFPR-MAP dataset to include other campus trajectories and conditions, and explore the scalability of Cognitive Sensor Sets to multi-robot systems.

This work advances the perception-action model with a cycle inspired by the human cognitive system using the Cognitive Blend Sensors method. Ultimately, the construction of the UFPR-MAP with ground-truth referenced by points obtained through geodetic and topographic

methods represents a significant contribution to the field of multimodal sensor fusion. It provides data obtained with a robot traversing between indoor and outdoor environments, gathering rich raw data from various types of standard sensors.

REFERÊNCIAS

- Alkhawaja, F., Jaradat, M., and Romdhane, L. (2019). Techniques of Indoor Positioning Systems (IPS): A Survey. In *2019 Advances in Science and Engineering Technology International Conferences (ASET)*, pages 1–8. IEEE.
- Anderson, J. R. (1995). *The Architecture of Cognition*. Psychology Press.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., and Qin, Y. (2004). An Integrated Theory of the Mind. *Psychological Review*, 111(4):1036–1060.
- Barricelli, B. R., Casiraghi, E., and Fogli, D. (2019). A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access*, 7(MI):167653–167671.
- Berto, L. M. (2020). *Exploring Cognitive Functions in Robotics*. PhD thesis, UNICAMP.
- Berto, L. M., de L. Rossi, L., Rohmer, E., Costa, P. D. P., Simoes, A. S., Gudwin, R. R., and Colombini, E. L. (2020). Learning over the Attentional Space with Mobile Robots. In *2020 Joint IEEE 10th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, pages 1–7. IEEE.
- Blanco-Claraco, J.-L., Moreno-Dueñas, F.-Á., and González-Jiménez, J. (2014). The Málaga urban dataset: High-rate stereo and LiDAR in a realistic urban scenario. *The International Journal of Robotics Research*, 33(2):207–214.
- Bosch Sensortec (2021). Datasheet BNO055: intelligent 9-axis absolute orientation sensor. Technical report.
- Brooks, R. R. and Iyengar, S. (1997). *Multi-Sensor Fusion: Fundamentals and Applications with Software*. Prentice Hall PTR.
- Carlevaris-Bianco, N., Ushani, A. K., and Eustice, R. M. (2016). University of Michigan North Campus long-term vision and lidar dataset. *The International Journal of Robotics Research*, 35(9):1023–1035.
- Carvalho, F., Santos, E., Iabrudi, A., Chaimowicz, L., and Campos, M. (2012). Indoor Wireless Sensor Localization using Mobile Robot and RSSI. In *2012 IEEE 9th International Conference on Mobile Ad-Hoc and Sensor Systems (MASS 2012)*, pages 1–6. IEEE.
- Chandrasekaran, B. and Conrad, J. M. (2016). Sensor fusion using a selective sensor framework to achieve decision and task execution. *Conference Proceedings - IEEE SOUTHEASTCON*, 2016-July:1–7.
- Chandrasekaran, B., Gangadhar, S., and Conrad, J. M. (2017). A survey of multisensor fusion techniques, architectures and methodologies. In *SoutheastCon 2017*, pages 1–8. IEEE.
- Chen, C., Rosa, S., Lu, C. X., Wang, B., Trigoni, N., and Markham, A. (2025). Learning Selective Sensor Fusion for State Estimation. *IEEE Transactions on Neural Networks and Learning Systems*, 36(3):4103–4117.

- Chen, C., Rosa, S., Miao, Y., Lu, C. X., Wu, W., Markham, A., and Trigoni, N. (2019). Selective Sensor Fusion for Neural Visual-Inertial Odometry. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2019-June, pages 10534–10543. IEEE.
- Chen, C., Wang, B., Lu, C. X., Trigoni, N., and Markham, A. (2020). A Survey on Deep Learning for Localization and Mapping: Towards the Age of Spatial Machine Intelligence. *arXiv*.
- Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., and Schiele, B. (2016). The Cityscapes Dataset for Semantic Urban Scene Understanding. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2016-Decem, pages 3213–3223. IEEE.
- de Souza, E. M. (2008). *Análise de wavelets para detecção e correção do multicaminho no posicionamento relativo GNSS estático e cinemático*. PhD thesis, UFPR.
- Drumheller, M. (1987). Mobile Robot Localization Using Sonar. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(2):325–332.
- Durrant-Whyte, H. F. (1988). Sensor Models and Multisensor Integration. *The International Journal of Robotics Research*, 7(6):97–113.
- ElecFreaks (2004). Datasheet - HCSR04 (Ultrasonic Ranging Module). Technical report.
- Everett, H. R. (1995). *Sensors for Mobile Robots - Theory and Application*. A K Peters/CRC Press, first edition.
- Fchollet (2021). Image classification from scratch. https://keras.io/examples/vision/image_classification_from_scratch/.
- Fuster, J. M. (2005). *Cortex and Mind: Unifying Cognition*. Oxford University Press.
- Geiger, A., Lenz, P., Stiller, C., and Urtasun, R. (2013). Vision meets robotics: The KITTI dataset. *The International Journal of Robotics Research*, (October):1–6.
- Glaessgen, E. and Stargel, D. (2012). The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles. In *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, Reston, Virigina. American Institute of Aeronautics and Astronautics.
- Gomes, A. (2023). *Avaliação e Mitigação do Efeito do Multicaminho no Posicionamento GNSS via Smartphones*. PhD thesis, UFPR.
- Grieves, M. (2014). Digital Twin : Manufacturing Excellence through Virtual Factory Replication. *White Paper*, (March):1–7.
- Guimarães, R. L., de Oliveira, A. S., Fabro, J. A., Becker, T., and Brenner, V. A. (2016). ROS Navigation: Concepts and Tutorial. In *Studies in Computational Intelligence*, volume 625, pages 121–160.
- Haykin, S. (2012). *Cognitive Dynamic Systems: Perception-action Cycle, Radar and Radio*. Cambridge University Press.

- HOKUYO (2024). Ust-30lx laser scanning rangefinder. <https://www.hokuyo-aut.jp/search/?cate01=1>.
- IniVation (2020). Understanding the Performance of Neuromorphic Event-based Vision Sensors. <https://inivation.com/wp-content/uploads/2020/05/White-Paper-May-2020.pdf>.
- Iovescu, C. and Rao, S. (2017). The fundamentals of millimeter wave sensors. *Texas Instruments*.
- Joseph, L. (2018). *Mastering ROS for Robotics Programming: Design, build, and simulate complex robots using the Robot Operating System*. Packt Publishing, second edition.
- Kam, M., Zhu, X., and Kalata, P. (1997). Sensor Fusion for Mobile Robot Navigation. *Proceedings of the IEEE*, 85(1):108–119.
- Khaleghi, B., Khamis, A., Karray, F. O., and Razavi, S. N. (2013). Multisensor data fusion: A review of the state-of-the-art. *Information Fusion*, 14(1):28–44.
- Kunemund, F., Lategahn, J., and Rohrig, C. (2009). WLAN mobile robot localization with sensor fusion. In *2009 IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, number September, pages 649–654. IEEE.
- Lange, D. (2019). Cognitive Robotics: Making Robots Sense, Understand, and Interact. *IEEE Computer Society*, 52(12):39–44.
- Lee, Y. K., Eu, K. S., and Choy, C. W. (2020). Simultaneous Localization and Mapping with Basic Cognitive Understanding of Environments. In *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, pages 1–6. IEEE.
- LEICK, A. (2004). *GPS Satellite Surveying*. John Wiley & Sons, third edition.
- Leonard, J. J. and Durrant-Whyte, H. F. (1992). *Directed Sonar Sensing for Mobile Robot Navigation*. Springer US, Boston, MA.
- Liu, H. and Darabi, H. (2007). Survey of Wireless Indoor Positioning Techniques and Systems. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, 37(6):1067–1080.
- Lowry, S., Sunderhauf, N., Newman, P., Leonard, J. J., Cox, D., Corke, P., and Milford, M. J. (2016). Visual Place Recognition: A Survey. *IEEE Transactions on Robotics*, 32(1):1–19.
- Luo, R. and Kay, M. (1989). Multisensor integration and fusion in intelligent systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(5):901–931.
- Maddern, W., Pascoe, G., Linegar, C., and Newman, P. (2017). 1 Year, 1000km: The Oxford RobotCar Dataset. *The International Journal of Robotics Research (IJRR)*, 36(1):3–15.
- Magrin, C. E., Brito, R. C., and Todt, E. (2019). A Systematic Mapping Study on Multi-Sensor Fusion in Wheeled Mobile Robot Self-Localization. In *2019 Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR) and 2019 Workshop on Robotics in Education (WRE)*, pages 132–137. IEEE.

- Magrin, C. E., Del Conte, G., and Todt, E. (2021). Creating a Digital Twin as an Open Source Learning Tool for Mobile Robotics. In *2021 Latin American Robotics Symposium (LARS), 2021 Brazilian Symposium on Robotics (SBR), and 2021 Workshop on Robotics in Education (WRE)*, pages 13–18. IEEE.
- Magrin, C. E. and Todt, E. (2019a). Multi-Sensor Fusion Method Based on Artificial Neural Network for Mobile Robot Self-Localization. In *2019 Latin American Robotics Symposium (LARS), 2019 Brazilian Symposium on Robotics (SBR) and 2019 Workshop on Robotics in Education (WRE)*, pages 138–143. IEEE.
- Magrin, C. E. and Todt, E. (2019b). Simulation of a Mobile Robot Localization based on Hierarchical Sensor Fusion. In *X Computer on the Beach*, pages 51–60, Florianopolis.
- Magrin, C. E. and Todt, E. (2019c). UFPR-RSFM Dataset (RSS Sonar Fingerprinting Map) V1.0. <https://web.inf.ufpr.br/vri/databases/ufpr-rsfm>.
- Magrin, C. E. and Todt, E. (2024). Multi-sensor mobile robots for sensor blend sets in indoor and outdoor environments. In *Proceedings of the V Brazilian Humanoid Robot Workshop (BRAHUR) and VI Brazilian Workshop on Service Robotics (BRASERO)*, page 6. Even3.
- Magrin, C. E. S. and Todt, E. (2016). Hierarchical Sensor Fusion Method Based on Fingerprint kNN and Fuzzy Features Weighting for Indoor Localization of a Mobile Robot Platform. In *2016 XIII Latin American Robotics Symposium and IV Brazilian Robotics Symposium (LARS/SBR)*, pages 305–310. IEEE.
- Malyavej, V. and Udomthanatheera, P. (2014). RSSI/IMU sensor fusion-based localization using unscented Kalman filter. In *The 20th Asia-Pacific Conference on Communication (APCC2014)*, pages 227–232. IEEE.
- Meyer, D. E. and Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part I. Basic mechanisms. *Psychological Review*, 104(1):3–65.
- Mitchell, H. (2007). *Multi-sensor data fusion*. Springer.
- Murphy, R. (2000). *Introduction to AI robotics*. A Bradford Book, 1st edition.
- Naghshvarianjahromi, M., Kumar, S., and Deen, M. J. (2020). Natural Brain-Inspired Intelligence for Non-Gaussian and Nonlinear Environments with Finite Memory. *Applied Sciences*, 10(3):1150.
- Newell, A. (1994). *Unified Theories of Cognition*. Harvard University Press; Reprint edition.
- OpenRobotics (2021). Why ros? it's the fastest way to build a robot! <https://www.ros.org/>.
- OpenRobotics (2024). Robot operating system provides libraries and tools to help software developers create robot applications. <http://wiki.ros.org/>.
- Pandey, G., McBride, J. R., and Eustice, R. M. (2011). Ford Campus vision and lidar data set. *The International Journal of Robotics Research*, 30(13):1543–1552.
- Paraense, A. L., Raizer, K., de Paula, S. M., Rohmer, E., and Gudwin, R. R. (2016). The cognitive systems toolkit and the CST reference cognitive architecture. *Biologically Inspired Cognitive Architectures*, 17:32–48.

- Pires, F., Cachada, A., Barbosa, J., Moreira, A. P., and Leitao, P. (2019). Digital Twin in Industry 4.0: Technologies, Applications and Challenges. In *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*, volume 2019-July, pages 721–726. IEEE.
- Qingqing, L., Xianjia, Y., Queralta, J. P., and Westerlund, T. (2022). Multi-Modal Lidar Dataset for Benchmarking General-Purpose Localization and Mapping Algorithms. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3837–3844. IEEE.
- Ramicic, M. and Bonarini, A. (2020). Selective Perception as a Mechanism to Adapt Agents to the Environment: An Evolutionary Approach. *IEEE Transactions on Cognitive and Developmental Systems*, 12(1):64–72.
- Rathore, H. and Bhadauria, V. (2022). Intelligent Decision Making in Autonomous Vehicles using Cognition Aided Reinforcement Learning. In *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, volume 2022-April, pages 524–529. IEEE.
- Rossi, L. d. L. (2021). *Aprendizado sensório-motor em robôs cognitivos utilizando modelo CST-CONAIM*. PhD thesis, UNESP.
- Santos, J. M., Portugal, D., and Rocha, R. P. (2013). An evaluation of 2D SLAM techniques available in Robot Operating System. In *2013 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pages 1–6. IEEE.
- Sarlin, P.-E., Debraine, F., Dymczyk, M., Siegwart, R., and Cadena, C. (2018). Leveraging Deep Visual Descriptors for Hierarchical Efficient Localization. *arXiv*, pages 1–10.
- Siegwart, R., Nourbakhsh, I. R., and Scaramuzza, D. (2011). *Introduction to Autonomous Mobile Robots*. MIT Press, second edition.
- Souza, E. M. (2004). Efeito de multicaminho de alta frequência no posicionamento relativo GPS estático: Detecção e atenuação utilizando Wavelets.
- STMicroelectronics (2024). Memos and sensors. <https://www.st.com/en/mems-and-sensors.html>.
- Stork, T. (2000). Electronic Compass Design using KMZ51 and KMZ52. *Philips Semiconductors - Application Note AN00022*, page 38.
- Sturm, J., Engelhard, N., Endres, F., Burgard, W., and Cremers, D. (2012). A benchmark for the evaluation of rgb-d slam systems. In *Proc. of the International Conference on Intelligent Robot Systems (IROS)*.
- SUCI, F. M. (2012). Estudo sobre a estabilidade da rede geodésica de monitoramento da UHE Salto Caxias.
- Sun, P., Kretschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., Tsui, P., Guo, J., Zhou, Y., Chai, Y., Caine, B., Vasudevan, V., Han, W., Ngiam, J., Zhao, H., Timofeev, A., Ettinger, S., Krivokon, M., Gao, A., Joshi, A., Zhang, Y., Shlens, J., Chen, Z., and Anguelov, D. (2020). Scalability in Perception for Autonomous Driving: Waymo Open Dataset. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2443–2451. IEEE.
- Tao, F., Zhang, H., Liu, A., and Nee, A. Y. C. (2019a). Digital Twin in Industry: State-of-the-Art. *IEEE Transactions on Industrial Informatics*, 15(4):2405–2415.

- Tao, F. and Zhang, M. (2017). Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing. *IEEE Access*, 5:20418–20427.
- Tao, F., Zhang, M., and Nee, A. (2019b). *Digital Twin Driven Smart Manufacturing*. Academic Press Inc.
- Tian, K. and Mirza, K. (2022). Sensor Fusion for Octagon - an Indoor and Outdoor Autonomous Mobile Robot. *SysCon 2022 - 16th Annual IEEE International Systems Conference, Proceedings*, pages 1–5.
- Todt, E. and Torras, C. (2004). Detecting salient cues through illumination-invariant color ratios. *Robotics and Autonomous Systems*, 48(2-3):111–130.
- Torteeke, P. and Chundi, X. (2014). Indoor Positioning Based on Wi-Fi Fingerprint Technique Using Fuzzy K-Nearest Neighbor. *Proceedings of 2014 11th International Bhurban Conference on Applied Sciences & Technology (IBCAST) Islamabad, Pakistan, 14th - 18th January, 2014*, pages 461–465.
- Tzafestas, S. G. (2014). *Introduction to Mobile Robot Control*. Elsevier.
- Veiga, L. A. K., Zanetti, M. A. Z., and Faggion, P. L. (2007). Fundamentos de Topografia. *Livro*, page 179.
- Vossiek, M., Wiebking, L., Gulden, P., Wiegardt, J., Hoffmann, C., and Heide, P. (2003). Wireless Local Positioning. *IEEE Microwave Magazine*, 4(December):77–86.
- Wettscheck, D. and Aha, D. W. (1995). Weighting Features. *First International Conference on Case-Based Reasoning (ICCBR-95)*. NCARAI TR AIC-95-026, page 12.
- Xiao, P., Shao, Z., Hao, S., Zhang, Z., Chai, X., Jiao, J., Li, Z., Wu, J., Sun, K., Jiang, K., Wang, Y., and Yang, D. (2021). PandaSet: Advanced Sensor Suite Dataset for Autonomous Driving. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, volume 2021-Sept, pages 3095–3101. IEEE.
- Yang, R. and Zhang, H. (2014). RSSI-Based Fingerprint Positioning System for Indoor Wireless Network. In *Intelligent Computing in Smart Grid and Electrical Vehicles. Volume 463 of the series Communications in Computer and Information Science*, pages 313–319.
- Yin, J., Li, A., Li, T., Yu, W., and Zou, D. (2022). M2DGR: A Multi-Sensor and Multi-Scenario SLAM Dataset for Ground Robots. *IEEE Robotics and Automation Letters*, 7(2):2266–2273.
- Zhang, H., Tang, H., and Yan, R. (2019). Multi-Sensor Fusion for A Brain-Inspired SLAM System. In *2019 5th International Conference on Control, Automation and Robotics (ICCAR)*, pages 619–623. IEEE.
- Zhang, Y., Li, X., and Amin, M. (2010). Principles and Techniques of RFID Positioning. In *RFID Systems*, pages 389–415. John Wiley & Sons, Ltd, Chichester, UK.
- Zhao, X. and Luo, Q. (2008). Survey on Robot Multi-sensor Information Fusion Technology. *2008 7th World Congress on Intelligent Control and Automation*, pages 5019–5023.
- Zou Yi, Ho Yeong Khing, Chua Chin Seng, and Zhou Xiao Wei (2000). Multi-ultrasonic sensor fusion for mobile robots. In *Proceedings of the IEEE Intelligent Vehicles Symposium 2000 (Cat. No.00TH8511)*, number Mi, pages 387–391. IEEE.