

# Integrating Multimodal Perception into Ground Mobile Robots

Ricardo B. Sousa<sup>\*†</sup>, Héber Miguel Sobreira<sup>†</sup>, João G. Martins<sup>\*†</sup>,  
Paulo G. Costa<sup>\*†</sup>, Manuel F. Silva<sup>†‡</sup>, and António Paulo Moreira<sup>\*†</sup>

<sup>\*</sup>Faculty of Engineering, University of Porto (FEUP), 4200-465 Porto, Portugal. {rbs,paco,amoreira}@fe.up.pt

<sup>†</sup>INESC TEC – Institute for Systems and Computer Engineering, Technology and Science, 4200-465 Porto, Portugal.

{heber.m.sobreira,joao.g.martins}@inesctec.pt

<sup>‡</sup>ISEP, Polytechnic of Porto, 4249-015 Porto, Portugal. mss@isep.ipp.pt

**Abstract**—Multimodal perception systems enhance the robustness and adaptability of autonomous mobile robots by integrating heterogeneous sensor modalities, improving long-term localisation and mapping in dynamic environments and human-robot interaction. Current mobile platforms often focus on specific sensor configurations and prioritise cost-effectiveness, possibly limiting the flexibility of the user to extend the original robots further. This paper presents a methodology to integrate multimodal perception into a ground mobile platform, incorporating wheel odometry, 2D laser scanners, 3D Light Detection and Ranging (LiDAR), and RGBD cameras. The methodology describes the electronics design to power devices, firmware, computation and networking architecture aspects, and mechanical mounting for the sensory system based on 3D printing, laser cutting, and bending metal sheet processes. Experiments demonstrate the usage of the revised platform in 2D and 3D localisation and mapping and pallet pocket estimation applications. All the documentation and designs are accessible in a public repository.

**Index Terms**—Light Detection and Ranging (LiDAR), mobile robot, multimodal perception, open-source, RGBD camera.

## I. INTRODUCTION

Multimodal perception systems pose unique challenges for autonomous mobile robots, requiring the real-time integration of asynchronous and heterogeneous data streams. These systems provide rich information, enabling robust interactions between humans and robots [1]. Another advantage of multimodal perception is enhancing the robustness of localisation and mapping algorithms, particularly in long-term applications with scene appearance changes [2].

As a result, integrating multimodal perception into ground mobile platforms improves the robustness of autonomous systems by combining diverse sensor modalities. This integration requires the consideration of several factors, including mechanical mounting, electronics design to power the sensors, and the computing architecture to gather and process sensor data. Currently, most commercially available and research-oriented mobile platforms focus on specific sensor models. These platforms typically prioritise low-cost solutions or have multimodal perception with a particular sensor configuration.

This work is co-financed by Component 5 – Capitalization and Business Innovation, integrated in the Resilience Dimension of the Recovery and Resilience Plan within the scope of the Recovery and Resilience Mechanism (MRR) of the European Union (EU), framed in the Next Generation EU, for the period 2021–2026, within project GreenAuto, with reference 54.

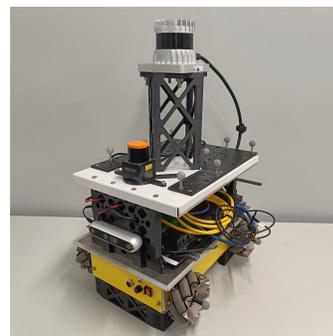


Fig. 1: Revised Hangfa Discovery Q2 mobile platform with multimodal perception (2D laser scanner – Hokuyo UST-10LX; 3D LiDAR – Ouster OS1 64; and RGBD camera – Intel RealSense D455).

This paper outlines a methodology to integrate multimodal perception into a ground mobile robot. The multimodal system incorporates 2D laser scanners, 3D Light Detection and Ranging (LiDAR), and RGBD (colour and depth) sensors. This system includes compatibility with at least four 2D lasers (Hokuyo UST-10LX, LD-19, RPLIDAR S2, and YDLIDAR X4), and several LiDARs (Livox Mid-360, Ouster OS, RoboSense Helios, Velodyne Puck/VLP) and RGBD camera (Intel RealSense D455/455f/456/L515, OAK-D and OAD-D Pro Series) models. The compatibility is achieved by relying on 3D printing and bent metal sheet fabrication processes with laser-cut hole patterns to standardise the sensor fixation onto the platform. Furthermore, this paper details a comprehensive overhaul of an omnidirectional platform, enabling wheel odometry, the computing unit’s integration onto the platform, and defining a network architecture for communication between sensors and other devices. Fig. 1 illustrates the revised Discovery Q2<sup>1</sup> platform used to demonstrate the application of the multimodal perception system. Experiments include tests on 2D and 3D Simultaneous Localisation and Mapping (SLAM) with 2D lasers and 3D LiDARs, and pallet pocket detection with 2D localisation and RGBD perception.

All electronics documentation, mechanical designs for the bent metal sheets, 3D-printed components, and other parts, along with the code developed in the scope of this paper,

<sup>1</sup><https://www.hangfa-europe.com/en/omni-robot/discovery>

are accessible in a public repository<sup>2</sup>. The repository aims to facilitate the replication of the adaptations made to the original Hangfa Discovery Q2 and multimodal perception integration into the robot, while also enabling the adaptation of this work to other ground mobile platforms and other applications.

## II. RELATED WORK

Mobile robots are gathering more interest due to their increasing applications across various domains, including warehouse logistics, healthcare, research, and education [3]. Consequently, commercial platforms have been developed to facilitate tests and research on autonomous navigation algorithms. One example is the iRobot Create 3<sup>3</sup> as an open-source educational robot designed for affordability and being compatible with the Robot Operating System (ROS) 2. Based on a Roomba vacuum cleaner, the differential-drive platform is equipped with sensors, such as InfraRed (IR) and cliff detectors, optical odometry, an Inertial Measurement Unit (IMU), and wheel encoders. The Clearpath TurtleBot 4<sup>4</sup> extends the functionality of the Create 3 by incorporating an RGB-D camera (OAK-D Pro) and a 2D laser scanner (RPLIDAR A1), enabling further spatial awareness of the robot's surroundings. However, the base Create 3 platform has limitations in terms of its relatively small battery (14.4 V, 1800 mAh) that powers the robot and the compute unit (NavQPlus on the original Create 3 and Raspberry Pi 4B on the TurtleBot 4 robot), while lacking native support for 3D LiDAR. Other commercial alternatives from Clearpath or Husarion<sup>5</sup> may provide autonomy packages supporting the integration of sensors like RGB-D cameras, IMUs, Global Navigation Satellite Systems (GNSS), and 3D LiDARs. Still, the sensors integration is typically closed-source and focus on specific sensor configurations.

Moreover, TraxBot [4] is a differential-track robot built upon the Traxster II educational kit and compatible with ROS. This platform focuses on affordability (300€, not considering the laptop), robustness (all hardware in aluminium or stainless steel), and operability indoors and outdoors. Although the authors mention compatibility with cameras and 2D lasers, no information is given on integrating those sensors into the platform. Another affordable open-source platform (£100) is the Mona [5] robot, designed for teaching and research purposes. A breakout board supporting the Teensy 3.2 with a WiFi module enables a ROS base station to receive IR sensors' data and send commands to the motors. Additionally, the platform includes two APDS-9960 RGB and gesture sensors to read colour data. The open-source Autonomous Mini Robot (AMiRo) [6] uses a custom operating system (AMiRO-OS), fully integrated with the hardware on the platform. AMiRo has sensors such as an accelerometer, gyroscope, and magnetometer and supports image processing with up to

four RGB cameras processed on a Xilinx Spartan 6 Field-Programmable Gate Array (FPGA). Nevertheless, TraxBot [4] and Mona [5] focus mainly on affordability, while all three, including AMiRo [6], do not integrate more complex perception systems, such as RGBD cameras or 2D/3D LiDARs.

Other works in the literature integrate RGBD cameras and 3D LiDARs into robotic platforms. The open-source DPoom [7] robot is a low-cost platform based on the TurtleBot3 Waffle Pi, equipped with the Intel RealSense D435i RGBD camera (also has an on-board IMU), and executes a real-time navigation algorithm on a LattePanda Alpha 864. Similarly, ROBOTONT [8] is an open-source omnidirectional robot with a polycarbonate chassis. The robot employs the RealSense D435i camera for visual and depth perception, with the low-level computation of the platform powered by an STM32 board and an Intel NUC i5 as the computation unit to run ROS nodes. The ROS-based Open-source Mobile Robot (ROMR) [9] is built from consumer hoverboard wheels on aluminium profiles, making all CAD designs available to the community. This platform integrates multiple sensors for perception, localisation, and mapping. Indeed, the sensory system includes the RealSense D435i for visual and depth perception, an MPU9250 IMU for tracking and localisation, and the Intel RealSense T265 and the RPLIDAR A2M8 for localisation and mapping. The ROMR platform uses an Arduino Mega with rosserial to connect the firmware with ROS, running the latter on an NVIDIA Jetson Nano. Moreover, Kim *et al.* [10] integrate multimodal perception on an Agile-X Hunter SE Ackerman robot for a comparative study of LiDAR SLAM. The extended platform incorporates a VectorNav VN-100 IMU, 3D LiDARs (Livox Mid-70 and Velodyne VLP-16), and a Real-Time Kinematics (RTK) GNSS receiver (Emlid Reach RS2), using aluminium profiles for the robot chassis. Kim *et al.* [10] also employ an RGBD camera (Intel RealSense D435), in order to colourise the point clouds obtained from the 3D LiDARs. Overall, while DPoom [7], ROBOTONT [8], and ROMR [9] are specific to the sensor configurations used in their original work, Kim *et al.* [10] is focused on comparing SLAM studies, not on how integrating multimodal sensors into the original Hunter SE platform.

Mobile robot competitions offer real-world platforms to test autonomous driving algorithms while integrating multimodal perception to extract information from the environment. AWS DeepRacer<sup>6</sup> is a fully autonomous 1/18th scale race car driven by reinforcement learning on a global racing league. The standard platform integrates a stereo camera and a 2D laser scanner. Furthermore, the F1TENTH Autonomous Vehicle System [11] is a versatile open-source competition platform for autonomous racing systems. The vehicle system, based on the Traxxas Slash 4x4 Premium Chassis, integrates a 2D laser scanner (Hokuyo UST-10LX), an optional RGBD camera (Intel RealSense D435i), and the NVIDIA Jetson Xavier NX as the computation unit. Still, robot competitions usually focus on the software development part, with the

<sup>2</sup>[https://gitlab.inesctec.pt/mrtd/open-source/inesctec\\_mrtd\\_hangfa\\_discovery\\_q2](https://gitlab.inesctec.pt/mrtd/open-source/inesctec_mrtd_hangfa_discovery_q2)

<sup>3</sup><https://edu.irobot.com/what-we-offer/create3>

<sup>4</sup><https://clearpathrobotics.com/turtlebot-4/>

<sup>5</sup><https://husarion.com/>

<sup>6</sup><https://aws.amazon.com/deepracer/>



Fig. 2: Original Hangfa Discovery Q2 mobile platform (A: recharger socket; B: power fuse; C: power switch).

hardware being mainly standard for all competitors.

### III. REVISED HANGFA DISCOVERY Q2 PLATFORM

The mobile platform used in this paper is the Hangfa Discovery Q2. First, the original robot (illustrated in Fig. 2) is introduced regarding its mechanical characteristics and electronics to show possible platform limitations. Next, an electronics redesign is proposed to enable further extendability of the platform for multimodal perception. The redesign focuses on increasing battery capacity and maximum current while improving the flexibility of the wheels' angular speed control and powering sensors and other external devices. Then, the firmware for the revised platform is developed to read the encoders for wheel odometry, control the motors' angular speed, and communicate with an external computing unit. Finally, a dedicated computing unit is added to the platform, and a networking architecture is proposed to communicate with sensors and allow remote access to the computer.

#### A. Original platform

The original Hangfa Discovery Q2 is a small robot platform ( $359 \times 313.5 \times 114$  mm and 7 kg weight, with a 23 mm ground clearance) developed by Hangfa Robotics<sup>7</sup>. This platform is a four-wheeled omnidirectional robot. Consequently, the robot has holonomic kinematics, allowing its linear and angular velocities to be decoupled from each other (i.e., drive in any direction without requiring rotation) [12]. The Discovery Q2 also has a coaxial pendulum suspension on its back wheels, enabling the four wheels to touch the ground while helping reduce some vibration when passing through rough ground.

Furthermore, the robot is equipped with the QMA10 mecanum wheels from Hangfa Robotics. These wheels are 101.6 mm in diameter and 45.7 mm wide, with a carbon steel hub, a 350 g weight, a 30 kg maximum load, and 10 rollers. Each roller has rubber and two bearings to fix them to the wheel, facilitating the platform to move smoothly and steadily. The shaft of an outer body bearing block bears the wheel in axial and radial load, where the motor shaft is only used to transfer torque, improving the robot's load capacity (rated at 20 kg). Moreover, the DC motors are the Faulhaber 2342 Original Equipment Manufacturer (OEM) model, with a 64:1 gear ratio and a 12 counts per revolution 5 V encoder with two channels at the motor shaft. These motors are rated with a 12 V voltage, 1.1 A current, 11 W output power, and 5800 Rotations Per Minute (RPM) speed, with a 6800 RPM no-load speed.

As a result, the robot has a 0.65 m/s maximum translational and a  $140^\circ/\text{s}$  maximum rotation speed.

Regarding the electronics, the original robot has a 12 V DC 10400 mAh Lithium-ion (Li-ion) battery. This battery allows more than 10 hours of autonomy, with a 3 kg load, a moving speed of 0.5 m/s, and a 70% moving rate. The charger is 100 ~ 240 V AC with a DC output rated at 12.6 V @ 3.0 A. Next, the IFB1205 board provides the 12 V and 5 V DC power buses and makes Controller Area Network (CAN) and RS232 interfaces accessible to other modules. The board also provides a 5 V @ 5 A DC output for external devices. Moreover, the robot has a main power switch to turn ON/OFF and a 10 A power fuse for electrical protection.

As for the motor drivers, the IMDR4 module drives the four motors and provides closed-loop control. This module implements the motion control algorithm of the four omnidirectional wheels, where the user can control the robot's linear and angular velocity or the individual speed of the wheels. The control is made through the CAN bus and RS232 interfaces. However, the IMDR4 module does not provide data for wheeled odometry to estimate the robot's pose through the robot's kinematic model and the displacement of the wheels.

The original platform provides a C# Software Development Kit (SDK) to communicate with the internal STM32F407 microcontroller. The latter is part of the RHF407 development board. The user may extend the platform by programming the RHF407 board, connecting devices to the CAN bus, or adding accessories such as a remote controller and laser sensors.

Still, the original platform has some drawbacks. First, the Discovery Q2 does not have a computing unit, e.g., to run ROS-based nodes. Next, wheeled odometry data is not available to the user. The encoder signal of the motors could be derived to another microcontroller, counting the pulses based on the quadrature of dual channel encoders [13]. However, this approach requires having two microcontrollers in the robot. Also, the IMDR4 board does not have an interface to change the internal closed-loop control. Furthermore, only a 5 V @ 5 A DC external output is provided to the user, limiting the possibilities of powering external 2D and 3D LiDARs, or other equipment that may require more current or a different voltage level. Lastly, the platform does not support natively 3D LiDAR or RGBD cameras.

#### B. Electronics redesign

So, this paper proposes a complete redesign of the Discovery Q2 platform's electronics. Fig. 3 presents the proposed electronics layout, which integrates custom 3D-printed components to accommodate all the electronic devices, simplifying the assembly and improving the organization. Moreover, the 3D-printed structure features a two-level design. The baseplate hosts the motor drivers (Cytron MD10C R3) and a passive Battery Management System (BMS) for 3S Lithium Polymer (LiPo) batteries with a maximum current of 20 A. Next, the upper level houses the microcontroller (Arduino Mega 2560 with a proto shield to host encoders and motor drivers' connectors) and a 5–30 V to 1.25–30 V DC/DC Buck-Boost

<sup>7</sup><https://www.hangfa-europe.com/>

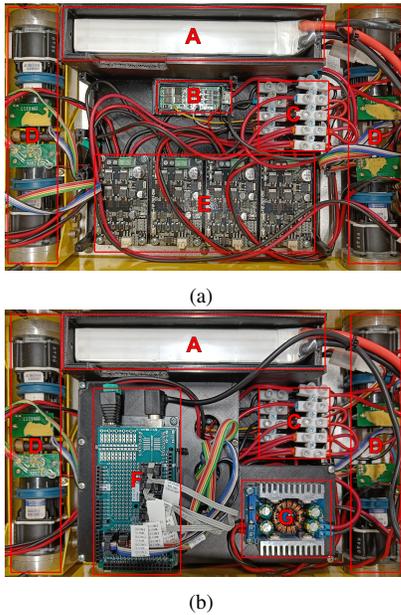


Fig. 3: Electronics redesign to integrate multimodal perception into the Discovery Q2 platform: (a) baseplate; (b) 1st level (A: battery; B: BMS; C: voltage and ground electric buses; D: wheel motors; E: motor drivers; F: microcontroller; G: DC/DC Buck-Boost converter).

Converter with a maximum current of 8 A. Then, the battery (Tattu 10000 mAh 11.1 V 3S LiPo with a 150 A maximum continuous discharge current) and the electrical connectors (voltage and ground electrical buses) are accessible to both the base and upper levels of the electronics structure.

In terms of powering the robot, the 3S battery output connector and its balance plug are directly connected to the cell interface of the BMS (0 V, 3.7 V, 7.4 V, and 11.1 V). The BMS's output is then connected to the robot's ON/OFF switch and a new 20 A power fuse, ensuring protected power delivery to all electronic components. This configuration enables the BMS to passively maintain the battery cells' balance during charging and discharging. For regulated voltage, the DC/DC Buck-Boost converter is connected to the BMS output. This converter can provide 12 V or 24 V DC-regulated power supply voltages, explicitly chosen for their compatibility with all LiDAR sensors considered for the platform (both in terms of voltage level and being regulated). Ouster OS sensors are suitable for 12 and 24 V DC nominal systems. The voltage range of the Livox Mid-360 (9–27 V) and the RoboSense Helios Series (9–32 V) is also compatible with 12 and 24 V DC. As for Velodyne Puck Series, its voltage range is 9–18 V DC, requiring to have the converter output to 12 V DC. Even if the users may have different requirements in terms of output voltage, the converter has a wide voltage adjustment range (1.25–30 V). Regarding recharging the platform, the BMS has a recharging voltage between 12.6 and 13.0 V and a maximum current of 10 A, making it compatible with the original charger of 12.6 V @ 3.0 A from the Discovery Q2 platform.

As for the motors, each one is connected to a single-channel motor driver MD10C R3. The platform's manual does not state

the maximum current of the OEM Faulhaber 2342 motor, only the 12 V and 1.1 A rated voltage and current, respectively. Nevertheless, the maximum 13 A continuous and 30 A peak (for 10 seconds) currents of MD10C R3 should be adequate to drive the motors. Next, the drivers are powered directly by the BMS, in which the motor's 12 V rated voltage aligns with the range of the 3S LiPo battery (approximately 9.6–12.6 V). Finally, the Pulse-Width Modulation (PWM) and direction control inputs of the drivers are linked to the Arduino Mega Proto Shield via 6-way Insulation-Displacement Contact (IDC) connectors, consistent with those used for the motors' encoders. This approach simplifies the interface for drivers and encoders by using a single connector type.

The microcontroller is also powered directly by the BMS output, as its input voltage range of 6–20 V is compatible with 3S LiPo batteries. Instead of relying on the Arduino Mega's internal pull-up resistors (20–50 k $\Omega$ ), external 3.3 k $\Omega$  pull-up resistors are soldered to the channels A and B inputs of the four encoders on the microcontroller's proto shield in order to reduce signal noise and improve the quadrature between channels A and B. These resistors are connected to the 5 V rail provided by the Arduino to form pull-up resistors. The same 5 V rail also powers the motors' encoders.

In summary, the electronics redesign proposed in this paper for the original Discovery Q2 introduces several improvements by enhancing its power capability and improving control flexibility. Indeed, the maximum current increases from 10 A to 20 A. Even though the official user manual does not clarify the maximum current supported by the original Li-ion battery, the 10 A value is based on the power fuse present previously in the platform. Still, the 20 A is only limited by the BMS (the newer battery supports a continuous discharge current up to 150 A), allowing the user to upgrade to a 40 A BMS or even a higher current capacity. This increase in power capability also enables additional external devices, such as 2D/3D LiDAR systems. Also, connecting encoders and PWM signals to the microcontroller allows users to implement different closed-loop control architectures, such as Proportional-Integrative (PI) controllers for angular speed control of the motors.

### C. Firmware

The firmware on the revised platform is implemented on the microcontroller Arduino Mega 2560. This paper builds upon the firmware developed for the Robot@Factory 4.0 competition [12]. The firmware implements the same methodology for processing the encoder signals based on the quadrature inherent to two-channel encoders. This approach accounts for the motor and the encoder's specifications, including the 64:1 gear ratio on the motors and the 12 counts per revolution resolution of the encoders present on the original robot. As a result, this configuration enables a resolution of 48 and 3072 pulses per revolution on the motors' shaft and at the wheels, respectively. The latter resolution corresponds to an angular resolution of 0.117°/pulse at the wheel shaft.

Similarly, the firmware incorporates the serial communication protocol based on the channels library used in the

previous work [12]. This protocol defines the data exchange between the microcontroller and an external computing unit. Furthermore, the channels protocol defines a single compact message type: the first byte specifies the data channel (type of information), and the subsequent data packet represents the data value. The latter is represented in binary (e.g., integer or float) or ASCII (hexadecimal representation of the data value as characters), corresponding to 4 and 8 bytes for the size of the data value on the package, respectively. Thus, the firmware follows a similar channels configuration from the previous work, communicating the pulses count of the wheels' encoders ( $g-j$ ), interval time between control cycles ( $k$ ), reset signal ( $r$ ), reference speed of the four motors ( $G-J$ ), and the PWM value applied to the motors if needed ( $K$ ).

Although this work adopts the same PI-based control system for the wheels' angular speed control from the previous work [12], this paper introduces improvements in PWM generation and code optimisations. The Robot@Factory framework uses the Adafruit Motor Shield v2 for the Arduino Mega to drive four motors. The shield has limitations in terms of maximum current and switching frequency, having a maximum of 1.2 A per motor (with a maximum 3 A peak for a duration of around 20 ms), close to the rated current of 1.1 A for the Faulhaber motors. Furthermore, the shield's reliance on an internal PWM driver chip and I2C bus communication for motor control results in a maximum PWM frequency supported by the shield's library<sup>8</sup> of approximately 1.6 kHz. This frequency is within the audible range, leading to undesired acoustic noises during operation.

Consequently, this paper adopts drivers that support external PWM signal generation (MD10C R3 supports up to 20 kHz switching frequency). The PWM is generated from the microcontroller, using *TimerOne* & *TimerThree*<sup>9</sup> libraries. As a result, the proposed firmware achieves 100 Hz for the closed-loop angular velocity speed control of the motors, instead of 50 Hz on the previous work [12], while not generating audible noise. This improvement on the firmware loop rate leads to the odometry data being available also at a 100 Hz.

#### D. Computing unit & networking

The computing unit selected for the revised Hangfa Discovery Q2 platform is the LattePanda 3 Delta Single Board Computer (SBC), illustrated in Fig. 4. This SBC has the Intel Celeron N5105 quad-core processor with four threads, a base clock of 2.00 GHz (boosting up to 2.90 GHz), and a Thermal Design Power (TDP) of 10 W. The processor includes an internal Intel UHD Graphics GPU with 24 execution units, operating at 450–800 MHz, and supports up to three external monitors. Moreover, the LattePanda 3 Delta features 8 GB of LPDDR4 RAM at 2933 MHz and 64 GB of onboard eMMC storage. The SBC also incorporates an M.2 PCIe 3.0 interface, supporting an NVMe SSD to increase storage capacity and performance. Indeed, in this paper, the Samsung 970 EVO Plus

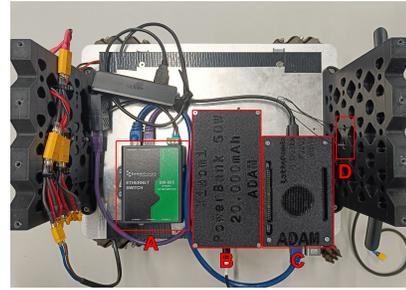


Fig. 4: LattePanda 3 Delta Single Board Computer (SBC) and the networking solution for the revised platform (A: Ethernet switch; B: computer battery; C: LattePanda 3 Delta; D: WiFi antennas).

250 GB NVMe SSD has been added as the primary storage device to improve both storage capacity and read/write speeds.

Regarding USB connectivity, the LattePanda 3 Delta provides two USB 3.2 Gen 1 ports, one USB 3.2 Gen 2 port (all Type-A), and a USB Type-C port with Power Delivery (PD), which can power the SBC. As for network connectivity, the SBC supports WiFi 6 at 2.4/5 GHz, Bluetooth 5.2, and 2.5 Gbps Ethernet. One advantage of the LattePanda 3 Delta is its x64-bit architecture, offering enhanced compatibility with ROS 1 and ROS 2 environments and open-source ROS-based packages, compared to ARM alternatives such as Raspberry Pi boards. Although the onboard ATmega 32U4 available in the SBC is not used as the microcontroller, another advantage of the LattePanda 3 Delta is having that microcontroller that could be used alongside the main processor.

In terms of powering the SBC, the LattePanda 3 Delta offers three options. The first is the LattePanda UPS Hat, which utilises three 18650 Li-ion cells and supports power delivery through Type-C PD and 5.5 mm DC interfaces. A key benefit of this solution is its support for the HID-UPS protocol, enabling the operating system to recognise it as a battery device. However, the overall battery capacity of the UPS Hat would be limited to the maximum capacity of one 18650 Li-ion cell, around 3600 mAh. The second option powers the SBC with 12 V DC through the JST PH2.0-4Pin connector. This option is compatible with the DC/DC converter integrated into the revised platform, provided its output is regulated to 12 V. However, relying on the converter may reduce the overall current available for external devices (LattePanda 3 Delta requires at least 2 A), restrain the converter's output voltage to 12 V, or even need an additional converter for external devices if another voltage is needed such as 24 V DC.

Thus, in this work, the SBC is powered using the Xiaomi Mi 50 W 20000 mAh power bank via the USB Type-C PD interface. A drawback is requiring an additional adapter to recharge the power bank. Nevertheless, using a power bank isolates the SBC electrically from other internal electronics and increases the battery capacity compared to the UPS Hat, depending on the power bank the user selects. The SBC accepts PD-compliant devices at 15 V @ 3 A or 12 V @ 3 A.

As for the networking architecture, the SBC provides wireless and Ethernet connectivity. This paper proposes to use

<sup>8</sup>[https://github.com/adafruit/Adafruit\\_Motor\\_Shield\\_V2\\_Library](https://github.com/adafruit/Adafruit_Motor_Shield_V2_Library)

<sup>9</sup>[https://www.pjrc.com/teensy/td\\_libs\\_TimerOne.html](https://www.pjrc.com/teensy/td_libs_TimerOne.html)

WiFi to connect the SBC to networks available at the robot’s deployment site. This connection is facilitated by attaching two Bingfu Dual Band WiFi Antennas to the robot’s frame (see Fig. 4) and connecting them to the SBC’s onboard Intel AX201 WiFi 6 wireless card. Then, the 2.5 Gbps Ethernet port provides a local network to connect sensors or other devices to the robot by having an Ethernet switch. The switches proposed for use in this paper are the Brainboxes 5-Port 10/100 Mbps SW-005 or 1 Gbps SW-015 models, both compatible with the revised platform regarding input power (5–30 V).

#### IV. MULTIMODAL PERCEPTION DESIGN & INTEGRATION

Integrating multimodal perception into the revised platform involves mechanical sensor mounting and software integration with the computing unit. First, the mechanical design proposed for sensor fixation is presented, detailing the approach of using a bent metal sheet and 3D-printed components. Then, the sensors are integrated into the platform by leveraging the ROS compatibility of the revised platform and the sensor drivers.

##### A. Mechanical design & sensors integration

The mechanical design for integrating multimodal perception into the revised Discovery Q2 aims to facilitate the integration of different sensors and enable the indexation of sensor placement through CAD designs. This indexation facilitates an initial estimation for the sensors’ extrinsic parameters. As a result, a laser-cut bent metal sheet with evenly distributed M4 holes is designed to host sensors such as 2D lasers, 3D LiDARs, and possibly RGBD cameras (see Fig. 5a). The metal sheet is made of Aluminum AW 1050 with 2 mm thickness. The bent shape on all sides increases the overall stiffness of the sheet. As for the holes, the laser cut process used in this work can achieve a 0.1–0.2 mm precision for cutting the whole sheet and its holes. This tolerance allows a precise indexation of the sensors compared to manually drilling the metal sheet.

Moreover, the M4 holes form a  $5 \times 5$  matrix with a 60 mm spacing, covering an area of  $240 \times 240$  mm. As a result, the outer dimensions of the final bent metal sheet are  $360 \times 270$  mm, compared to the Discovery Q2’s dimensions of  $359 \times 313.5$  mm, avoiding exceeding significantly the base footprint of the robot. In order to secure the plate to the robot’s frame, eight M5 holes are positioned on the top surface of the metal sheet (four at the front and four at the back). These holes fixate the sheet to the 3D-printed PETG supports shown in Fig. 1 and Fig. 4. These supports are specifically designed for the Discovery Q2 platform to align with the four screws on the original top plate (see Fig. 2). The alignment provides the indexation of the bent metal sheet with respect to the base.

In terms of sensors placement, the ranging sensors (2D/3D LiDARs) are mounted on the bent metal sheet. This placement design minimises obstructions caused by the robot’s body. As illustrated in Fig. 1 and Fig. 5, 2D laser scanners are positioned near the front of the metal sheet. 3D LiDARs are mounted at the centre in an elevated position to avoid the sensor beams being obstructed by the robot’s body. Although a  $360^\circ$  2D laser scanner may have obstructions using the

TABLE I: Specification Comparison of the RGBD Sensors Intel RealSense D455, Intel RealSense L515, and OAK-D Pro

Specs		RS D455	RS L515	OAK-D Pro
Color	H/VFOV ( $^\circ$ )	90°/65°	70°/43°	69°/55°
	Res. (px)	1280 × 800	1920 × 1080	4056 × 3040
	Rate (fps)	30	30	60
Depth	Type	active stereo	LiDAR	active stereo
	H/VFOV ( $^\circ$ )	87°/58°	70°/55°	80°/55°
	Res. (px)	1280x720	1024x768	1280x800
	Rate (fps)	90	30	120
	Range (m)	0.6–6	0.25–9	0.7–12

TABLE II: Specification Comparison of the 2D Laser Scanners Hokuyo UST-10LX, LD-19, RPLIDAR S2, and YDLIDAR X4

Specs	UST-10LX	LD-19	RPLIDAR S2	X4
FOV/Res. ( $^\circ$ )	270°/0.125°	360°/0.8°	360°/0.1125°	360°/0.432–0.864°
Type	ToF	ToF	ToF	Triangulation
Rate (Hz)	40	~ 10	~ 10	~ 10
Range (m)	0.06–10	0.02–12	0.05–30	0.12–10

TABLE III: Specification Comparison of the 3D LiDARs Livox Mid-360, Ouster OS1 64 Rev C, RoboSense RS-HELIOS-5515, and Velodyne VLP-16

Specs	Mid-360	OS1 64	HELIOS-5515	VLP-16
H/VFOV ( $^\circ$ )	360°/59°	360°/45°	360°/70°	360°/20°
HRes. ( $^\circ$ )	–	0.176/0.352/0.703°	0.1/0.2/0.4°	0.1–0.4°
VRes. ( $^\circ$ )	–	0.71°	$\leq 1.33^\circ$	1.33°
VRange ( $^\circ$ )	-7+52°	-22.5+22.5°	-55+15°	-10+10°
VType	non-uniform	uniform	non-uniform	uniform
#channels	–	64	32	16
Rate (Hz)	10	10/20	5/10/20	5–20
Range (m)	0.1–70	0.3–120	0.2–150	100

placement proposed in this paper, those obstructions are due to the 3D LiDAR fixation support and not to the robot’s body. Also, the 3D LiDAR support may be dismantled when not needed or even retrieve a  $360^\circ$  2D point cloud from the LiDAR itself. RGBD cameras are placed on the front of the 3D-printed support that fixates the bent metal sheet to the robot (see Fig. 5b). This placement ensures an unobstructed Field Of View (FOV) to the front of the robot while avoiding stacking sensors on top of each other, facilitating the design of the fixation supports. Nevertheless, the platform users may place the sensors differently on the metal sheet, leveraging the holes pattern to index the sensors to the CAD designs.

Tables I–III present specification comparisons of the sensors considered in this paper for multimodal perception integration into the revised platform. The three RGBD cameras (Intel RealSense D455/L515, and OAK-D Pro) do not require separated DC power, only need an USB connection to the SBC. Furthermore, three 2D laser scanners (LD-19, RPLIDAR S2, and YDLIDAR X4) are also powered directly through the USB connection. In contrast, the Hokuyo UST-10LX sensor is connected to the DC/DC converter and the data transmission is provided through Ethernet. As for 3D LiDARs, similar to the Hokuyo sensor, the four LiDARs (Livox Mid-360, Ouster OS1 64 Rev C, and RoboSense RS-HELIOS-5515) are connected to the DC/DC converter and the Ethernet switch.

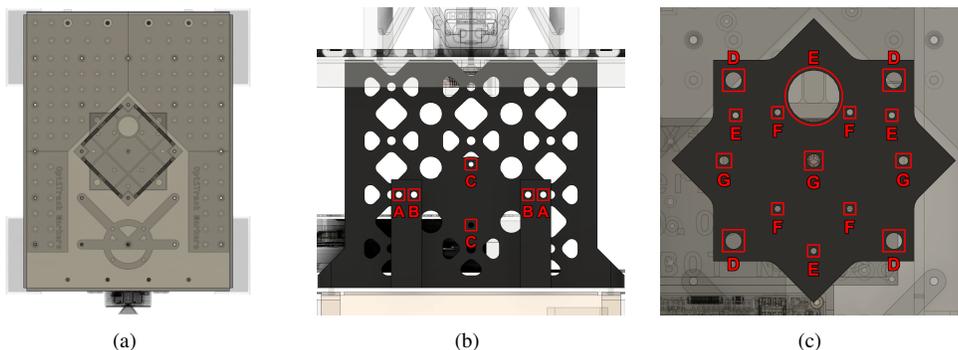


Fig. 5: 3D model views on AutoDesk Fusion 360: (a) top view of the bend metal sheet; (b) front support of the metal sheet for RGBD cameras (A: Intel RealSense D455/455f/456; B: OAK-D and OAD-D Pro Series; C: Intel RealSense L515); (c) support on top of the metal sheet for LiDARs 3D (D: Ouster OS sensors; E: RoboSense Helios Series; F: Livox Mid-360; G: Velodyne Puck/VLP).

Overall, the RGB cameras, 2D lasers, and LiDARs considered in this paper offer diversity in terms of sensor resolution, depth and ranging estimation types – active stereoscopy versus LiDAR on RGBD cameras, Time-of-Flight (ToF) versus triangulation on 2D lasers –, FOV resolution and range (including uniform versus non-uniform vertical FOVs on 3D LiDARs), and acquisition rate. This diversity allows the test of 2D/3D SLAM and object detection algorithms, among others, with different multimodal characteristics while having initial estimation for the sensors’ extrinsic parameters. Still, other sensors may be used as long as their power and communication requirements are compatible with the platform, only requiring adapting or even design new 3D-printed fixation supports.

### B. ROS integration

A ROS driver was developed for the revised platform based on a USB serial connection for the SBC to communicate with the firmware running on the microcontroller. The driver publishes the data from the wheel encoders and subscribes to the reference angular speed for the motors. Then, a ROS node computes the wheel odometry based on the kinematics of a four-wheeled omnidirectional robot [14]. Both the ROS driver and the odometry estimation node are compatible with ROS 1 and ROS 2 to improve the platform’s usability.

Since all the sensors considered in this work have ROS-compatible drivers, the data can be published directly in ROS. Integrating the sensors and the revised platform in ROS enables the data acquisition throughout the environment and execution of algorithms that leverage multimodal perception capabilities, either open-source or developed by the user.

## V. EXPERIMENTAL TESTS

The experimental evaluation of the proposed multimodal perception integration into a ground mobile platform focused on testing various robot perception applications. Fig. 6 presents the experimental results with the revised Discovery Q2 platform. First, the platform executes online 2D SLAM using wheel odometry and 2D laser with the SLAM Toolbox algorithm [15]. With the latter configured to skip only one scan, publishing the robot’s pose at 50 Hz, and considering

the Hokuyo UST-10LX’s rate of 40 Hz, the onboard SBC on the platform performed 2D SLAM without throttling. The second experiment executes 3D SLAM using wheel odometry and 3D LiDAR data with the VineSLAM [16] stack configured to 300 particles. Using the Livox Mid-360 sensor, the onboard computer achieved an online SLAM processing rate of approximately 7 Hz with a sensor rate of 10 Hz. When using the RoboSense RS-HELIOS-5515, the online execution rate decreases to approximately 3 Hz, possibly due to this sensor’s higher data output than the Mid-360 (576000 pts/s in single return mode versus 200000 pts/s, respectively).

As for the third test, a multimodal perception application is tested offline for pallet pocket estimation [17] with 2D localisation [18]. The data recording subscribed to all topics, including the ones from the 2D localisation system and the Intel RealSense L515 (set to  $640 \times 480$  px resolution and 6 Hz and 30 Hz for RGB and depth data, respectively). The offline processing demonstrates another application of the proposed multimodal system in this paper: gathering multimodal data and processing it offline for research and development. Still, a possible alternative to run the algorithm online could be using the LattePanda Sigma SBC (12-core, 16-thread Intel Core i5-1340P processor and 32 GB of RAM). This SBC is also compatible with power bank charging. However, the power bank must be upgraded to support at least 90 W @ 20 V. Videos from the experiments presented in this paper and additional tests are available online<sup>10</sup>.

## VI. CONCLUSIONS AND FUTURE WORK

In conclusion, this paper proposes a comprehensive methodology to integrate multimodal perception into a ground mobile robot leveraging 3D printing, laser cutting, and bending metal sheet fabrication processes. The methodology includes mechanical, electronics, firmware, computation and networking architecture aspects while providing an initial estimation of the extrinsic sensor parameters from CAD designs. While the integration focused on adapting the Hangfa Discovery Q2 platform for multimodal perception, the revision made to the

<sup>10</sup><https://www.youtube.com/playlist?list=PLvp8fJUEPxySkKsOrCN5FzjuhhSfVgSuR>

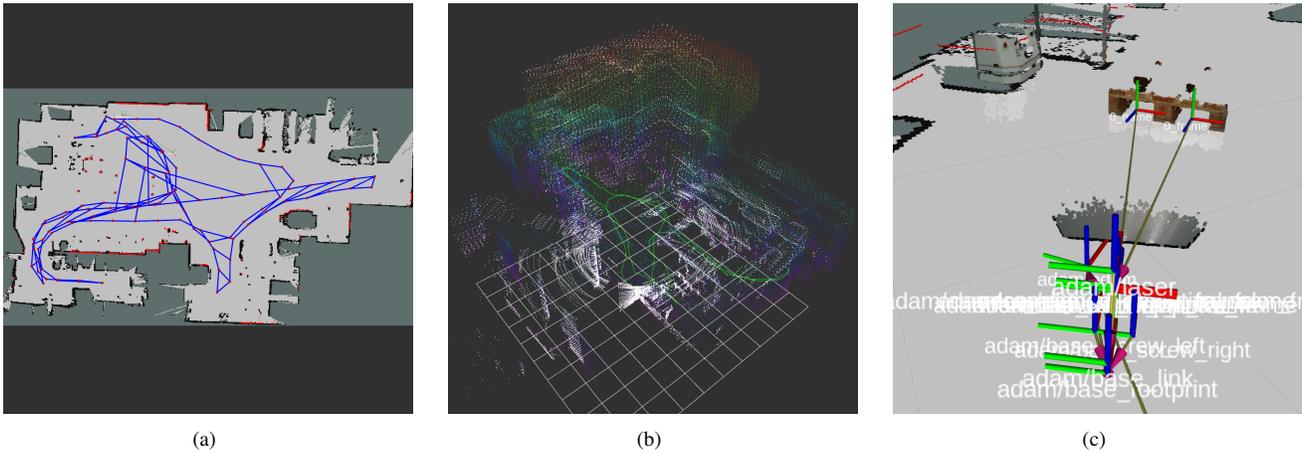


Fig. 6: Experimental tests of the revised platform with multimodal perception: (a) 2D SLAM [15] with wheel odometry and 2D laser scanner (Hokuyo UST-10LX); (b) 3D SLAM [16] with wheel odometry and 3D LiDAR (Livox Mid-360); (c) pallet pocket detection [17] with 2D localisation [18] and RGBD perception (Intel RealSense L515).

original robot can be extended to similar platforms or even adjusted to meet voltage, power, computation, and sensorisation requirements. The multimodal perception system is demonstrated in real-world experiments through online and offline data processing, showcasing its capability in applications such as 2D SLAM, 3D SLAM, and pallet pocket detection. All electronics documentation, mechanical designs, along with the code developed in the scope of this paper, are open-source and made available in a public repository<sup>2</sup>. For future work, the platform will be used and tested further for research and development, SLAM benchmarking, sensor calibration algorithms, and RGBD perception in industrial applications.

#### ACKNOWLEDGEMENTS

The authors would like to acknowledge all the support from the 5dpo Robotics Team<sup>11</sup>, Amorins & Silva<sup>12</sup> for their expertise in machining bent metal sheets, Hangfa Robotics<sup>7</sup> for their support related to the Discovery Q2 platform, and the LattePanda Team<sup>13</sup> for their sponsorship and technical support.

#### REFERENCES

- [1] J. Duncan, F. Alambeigi, and M. Pryor, "A survey of multimodal perception methods for human-robot interaction in social environments," *J. Hum.-Robot Interact.*, vol. 13, no. 4, p. 46, 2024.
- [2] R. Sousa, H. Sobreira, and A. Moreira, "A systematic literature review on long-term localization and mapping for mobile robots," *Journal of Field Robotics*, vol. 40, no. 5, pp. 1245–1322, 2023.
- [3] M. Alatise and G. Hancke, "A review on challenges of autonomous mobile robot and sensor fusion methods," *IEEE Access*, vol. 8, pp. 39 830–39 846, 2020.
- [4] A. Araújo, D. Portugal, M. Couceiro, C. Figueiredo, and R. Rocha, "TraxBot – assembling and programming of a mobile robotic platform," in *Proceedings of the 4th International Conference on Agents and Artificial Intelligence – Volume 2: ICAART*, 2012, pp. 301–304.
- [5] F. Arvin *et al.*, "Mona: an affordable open-source mobile robot for education and research," *Journal of Intelligent & Robotic Systems*, vol. 94, no. 3, pp. 761–775, 2019.
- [6] S. Herbrechtsmeier, T. Korthals, T. Schopping, and U. Rückert, "AMiRo: A modular & customizable open-source mini robot platform," in *2016 20th International Conference on System Theory, Control and Computing (ICSTCC)*, 2016, pp. 687–692.
- [7] T. Kim, S. Lim, G. Shin, G. Sim, and D. Yun, "An open-source low-cost mobile robot system with an RGB-D camera and efficient real-time navigation algorithm," *IEEE Access*, vol. 10, pp. 127 871–127 881, 2022.
- [8] R. Raudmäe *et al.*, "ROBOTONT – open-source and ROS-supported omnidirectional mobile robot for education and research," *HardwareX*, vol. 14, p. e00436, 2023.
- [9] L. Nwankwo, C. Fritze, K. Bartsch, and E. Rueckert, "ROMR: a ROS-based open-source mobile robot," vol. 14, p. e00426, 2023.
- [10] M. Kim, M. Zhou, S. Lee, and H. Lee, "Development of an autonomous mobile robot in the outdoor environments with a comparative survey of LiDAR SLAM," in *2022 22nd International Conference on Control, Automation and Systems (ICCAS)*, 2022, pp. 1990–1995.
- [11] M. O’Kelly, H. Zheng, D. Karthik, and R. Mangharam, "F1TENTH: an open-source evaluation environment for continuous control and reinforcement learning," in *Proceedings of the NeurIPS 2019 Competition and Demonstration Track*, vol. 123, 2020, pp. 77–89.
- [12] R. Sousa *et al.*, "A robotic framework for the Robot@Factory 4.0 competition," in *2024 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, 2024, pp. 66–73.
- [13] F. Sá, "Navigation system for omnidirectional mobile platform," Master’s thesis, Faculty of Engineering, University of Porto (FEUP), 2016. [Online]. Available: <https://hdl.handle.net/10216/85263>
- [14] R. Sousa, M. Petry, P. Costa, and A. Moreira, "OptiOdom: a generic approach for odometry calibration of wheeled mobile robots," *Journal of Intelligent & Robotic Systems*, vol. 105, no. 2, p. 39, 2022.
- [15] S. Macenski and I. Jambrecic, "SLAM toolbox: SLAM for the dynamic world," *Journal of Open Source Software*, vol. 6, no. 61, p. 2783, 2021.
- [16] A. Aguiar, F. Santos, L. Santos, A. Sousa, and J. Boaventura-Cunha, "Topological map-based approach for localization and mapping memory optimization," *Journal of Field Robotics*, vol. 40, no. 3, pp. 447–466, 2023.
- [17] D. Caldana *et al.*, "Pallet and pocket detection based on deep learning techniques," in *2024 7th Iberian Robotics Conference (ROBOT)*, 2024, pp. 1–8.
- [18] H. Sobreira *et al.*, "Map-matching algorithms for robot self-localization: a comparison between perfect match, iterative closest point and normal distributions transform," *Journal of Intelligent & Robotic Systems*, vol. 93, no. 3, pp. 533–546, 2019.

<sup>11</sup><https://5dpo.github.io/>

<sup>12</sup><https://amorinsesilva.pt/>

<sup>13</sup><https://www.lattepanda.com/>