

LITERATURE REVIEW

This chapter presents a review of existing literature on human-robot interaction (HRI) and the integration of augmented reality (AR) into mobile robotics. It discusses key advances, methodologies, and technologies relevant to this project.

2.1 INTRODUCTION TO HUMAN-ROBOT INTERACTION (HRI) AND AUGMENTED REALITY (AR)

Human-Robot Interaction (HRI) has seen substantial growth in both research and industrial applications, particularly with the advent of collaborative robots (cobots) that work alongside humans in dynamic environments. These robots, unlike their predecessors, are designed to safely interact with humans in shared workspaces, marking a significant shift in industrial automation. According to [7], HRI evolved from isolated robot systems in manufacturing to collaborative systems where robots assist humans in complex tasks. This shift is largely due to advancements in sensing, control, and human-machine interface technologies that allow robots to perceive their environment and interact intelligently.



Figure 2.1: Screen Display of work environment with AR additions

Augmented Reality (AR) has emerged as a powerful tool in enhancing HRI by providing a more intuitive and efficient way for humans to interact with robots. AR serves as a bridge that

overlays digital information onto the physical environment, enabling users to better understand the robot's actions and the task at hand [8].

2.1.1 Augmented Reality as a Tool in HRI

AR enhances communication between humans and robots by offering real-time visualizations and feedback, thus improving situational awareness and reducing errors in task execution. Suzuki et al. [8] propose a taxonomy of AR-enhanced HRI, highlighting the key areas where AR plays a role, such as task guidance, real-time interaction, and environment mapping. In particular, AR's ability to provide visual feedback significantly enhances the user experience by making robot operations more transparent and reducing cognitive load.

Moreover, AR facilitates dynamic task interaction by enabling users to provide real-time instructions to the robot. This is particularly beneficial in industrial applications where task conditions change frequently. AR not only improves task accuracy but also contributes to a more efficient workflow, as robots can quickly adjust their behavior based on user commands and visual markers such as ArUco codes [8].

2.1.2 Applications of AR in HRI

AR's applications in HRI span various industries, from manufacturing to healthcare. In manufacturing, AR has been utilized to guide robot operations through real-time visual overlays, improving precision and safety in tasks that require close human-robot collaboration [7]. Similarly, in healthcare, AR assists medical robots in performing delicate procedures by providing real-time feedback and guidance, enhancing both safety and operational efficiency [8].

Fiducial markers, such as ArUco, have become essential in improving robots' context-awareness. These markers, when integrated with AR systems, allow robots to perceive their surroundings more accurately, which enhances their ability to navigate and perform tasks autonomously [7].

2.1.3 Challenges and Future Directions

While AR has shown great promise in HRI, there are still challenges to overcome, particularly in real-time processing and system robustness in industrial settings. As noted by Suzuki et al. [8], further advancements in AR hardware and software are needed to make these systems more reliable and responsive in complex environments. Future research is expected to focus on improving AR's real-time capabilities and exploring new forms of interaction that can enhance

human-robot collaboration in dynamic settings [7].

2.2 MOBILE ROBOTICS AND AR INTEGRATION

The integration of Augmented Reality (AR) in mobile robotics has transformed how robots communicate their intentions and interact with their environment. AR not only enhances human-robot collaboration but also provides an intuitive interface for users to understand robot behavior in dynamic and complex environments. Several key studies have explored the use of AR to improve robot navigation, task execution, and real-time interaction between humans and robots.

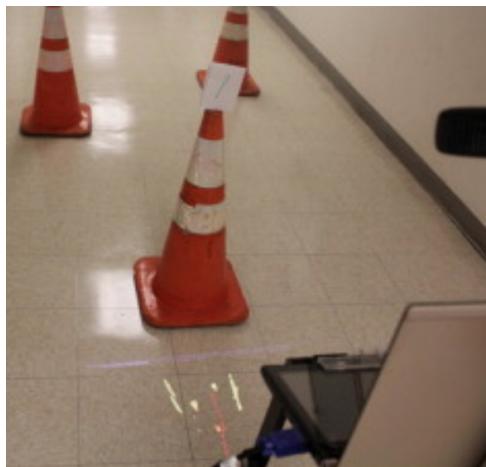
2.2.1 AR as a Communication Tool in Mobile Robotics

One of the primary challenges in human-robot interaction is communicating the robot's intended actions to human operators. Spatial Augmented Reality (SAR) has proven to be an effective method for achieving this. In the study by Green et al. [9], SAR is used to project information about the robot's intended movement directly onto the physical environment. This method allows users to better anticipate and respond to the robot's actions, reducing uncertainty and enhancing collaboration. Michalos et al. [10] demonstrated the application of AR in industrial settings, where visual feedback aids in task coordination between humans and robots, leading to improved task accuracy and faster response times.

Additionally, Covert et al. [11] explored how spatial AR techniques can improve communication between robots and humans. In their experiment, a robot used visual projections of arrows and simplified maps on the floor to indicate its short-, mid-, and long-term movement intentions. The study found that participants were able to predict the robot's movements with high confidence when the robot projected its intended path. This is especially useful in environments where humans and robots share a workspace, such as hospitals, museums, and factories.

2.2.2 Multimodal AR Interfaces for Human-Robot Interaction

In addition to visual feedback, multimodal interfaces that combine AR with other modalities, such as voice or gesture recognition, have further improved human-robot interaction. Green et al. [9] developed a multimodal AR interface for mobile robots, allowing users to interact with the robot through visual overlays, voice commands, and gesture-based inputs. This interface improved the efficiency and accuracy of task execution in collaborative environments. The study demonstrated that multimodal AR interfaces allow for more flexible and adaptive human-robot communication, making the interaction process more intuitive and reducing cognitive



(a) Robot estimating its intended movement using AR arrows [11].



(b) Robot estimating its intended movement using AR arrows [11].

load on the user.

2.2.3 Applications of AR in Mobile Robotics

AR's integration into mobile robotics is particularly useful in environments that require constant adaptation, such as manufacturing or logistics. Michalos et al. [10] applied AR in human-robot cooperation in industrial settings, where robots must adapt to rapidly changing conditions. AR provided real-time feedback to users, helping them coordinate with robots for tasks such as assembly or quality control. Moreover, AR visualizations helped robots better navigate dynamic environments, improving both safety and efficiency.

Furthermore, AR's ability to create real-time visual overlays in the robot's workspace has significant implications for task planning and execution. In particular, SAR-based interfaces allow users to visually instruct robots by marking control zones or highlighting obstacles, which the robot then interprets for navigation and task completion [11].

2.2.4 Challenges and Future Directions

While AR's role in mobile robotics has demonstrated significant potential, challenges remain in real-time processing, accurate object recognition, and user feedback latency. As Michalos et al. [10] noted, the integration of AR into robotic systems requires further advancements in sensor technology and computational power to ensure real-time responsiveness in dynamic environments. Future research will likely focus on improving AR's scalability in complex settings, refining multimodal interfaces, and enhancing the autonomy of mobile robots to handle more sophisticated tasks.

2.3 FIDUCIAL MARKER SYSTEMS IN HRI

2.3.1 Introduction to AR-Enhanced Interactions

Augmented Reality (AR) is increasingly used in robotics to facilitate different types of Human-Robot Interactions (HRI). These interactions include:

- **Visual Interaction:** AR overlays visual cues such as control zones, instructions, and real-time status feedback, enhancing human comprehension of the robot's operational environment.
- **Gesture-Based Interaction:** Human gestures can be interpreted by robots through AR-enhanced vision systems, reducing reliance on manual interfaces.
- **Speech and Command Interaction:** AR, combined with voice recognition, allows users to issue commands to robots with visual aids providing additional context.
- **Fiducial Marker-Based Interaction:** Fiducial markers such as ArUco or AprilTags provide robots with a way to interpret their environment, helping with localization, navigation, and object detection.

2.3.2 Fiducial Marker Systems Overview

In AR-enhanced HRI, fiducial marker systems are essential for real-time contextual interactions. These systems use visually distinct markers that robots detect through cameras. The most commonly used marker systems include ARToolKit, AprilTags, and ArUco markers, each with unique characteristics suited to various applications.

Table 2.1: Comparison of Fiducial Marker Systems

Marker System	Robustness	Detection Accuracy	Computational Efficiency	Efficiency	Applications
ARToolKit AprilTags ArUco Markers	Moderate High High	Good Excellent Good	Moderate Moderate-High High		Early AR applications Robotics Real-time robot navigation

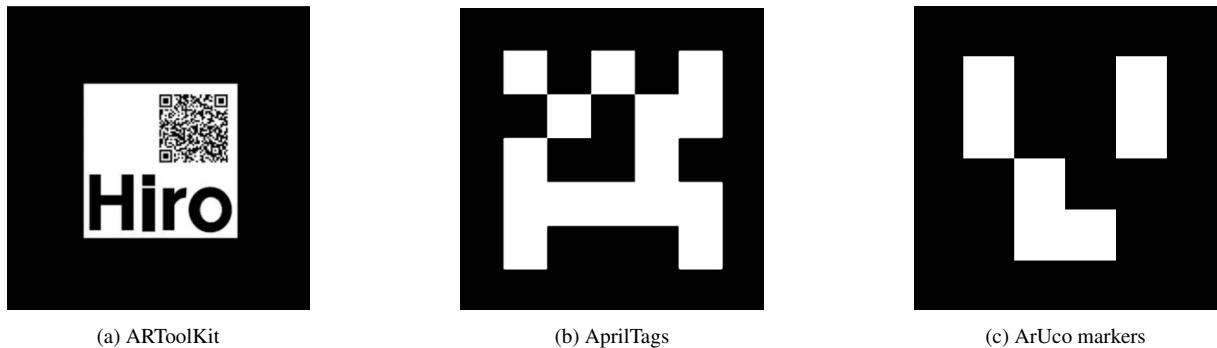


Figure 2.3: Comparison of different marker systems

2.3.2.1 *ARToolKit*

ARToolKit is one of the earliest fiducial marker systems developed for augmented reality applications. It was widely adopted in early AR systems due to its simplicity and moderate computational requirements. In HRI, ARToolKit has been used in early experiments involving robot navigation and control. However, its robustness is limited compared to more modern systems, and it struggles under challenging lighting conditions [8].

2.3.2.2 *AprilTags*

AprilTags are an improvement over ARToolKit, offering significantly higher robustness and detection accuracy. As [12] highlights, AprilTags are designed to provide accurate detection even in adverse environments with variable lighting and occlusion. AprilTags have been extensively used in industrial applications where precise localization is critical, such as warehouse automation and autonomous guided vehicles (AGVs). The system is computationally heavier than ARToolKit but provides superior reliability, making it suitable for applications that demand accuracy over speed [13].

2.3.2.3 *ArUco Markers*

ArUco markers are designed with computational efficiency in mind, offering high detection speed without sacrificing much in terms of accuracy. This makes them ideal for real-time applications, particularly in mobile robotics where swift responses are critical. [14] notes that ArUco markers have been used in various applications including navigation and task execution in dynamic environments such as warehouses. Their simplicity and widespread software support make them a popular choice for developers working on real-time robotic systems.

2.3.3 Choosing the Best Marker System for the Project

Given the requirements of the project, which involve real-time processing, dynamic navigation, and user interaction, ArUco markers are the most suitable choice. As [13] demonstrates, ArUco markers provide a good balance between computational efficiency and detection accuracy, making them ideal for mobile robotic systems that require fast decision-making. Additionally, the ease of integration with commonly used robotics software such as ROS (Robot Operating System) further justifies their selection for this project.

2.3.4 Visual Interaction in AR-Enhanced HRI

2.3.4.1 *Overview of Visual Interaction*

Visual interaction is one of the most common and effective forms of interaction in Augmented Reality (AR)-enhanced Human-Robot Interaction (HRI). It enables users to interact with robots by visualizing contextual information directly within their physical environment. This can include task instructions, control zones, and real-time feedback on the robot's status and performance.

In AR-enhanced visual interaction, the robot's actions and intentions are augmented with overlays that help the human user make informed decisions during interaction. [8] provides a comprehensive overview of how visual interaction enhances HRI by improving situational awareness and reducing the complexity of controlling robotic systems. This is especially useful in industrial applications, where workers interact with robots to execute precise tasks, such as assembly and quality control.

2.3.4.2 *Applications of Visual Interaction*

Visual interaction in AR has broad applicability across various fields, with significant benefits in improving task accuracy, communication, and overall efficiency. Some notable applications include:

- **Manufacturing and Assembly:** In industries such as automotive and aerospace, visual interaction is used to overlay task instructions on the assembly line, helping workers follow complex procedures. This reduces human error and increases productivity by providing step-by-step visual guidance [12].
- **Healthcare:** In robotic surgery and rehabilitation, visual interaction allows surgeons and

therapists to see real-time data overlaid on patients, enhancing precision and safety. AR enhances the accuracy of procedures by showing virtual boundaries and guidelines, ensuring that robots move correctly within the workspace [14].

- **Warehouse and Logistics:** In logistics and warehouse management, AR systems provide real-time visual cues to workers handling robotic systems for item sorting, picking, and delivery. [14] explores the use of AR to improve warehouse efficiency by overlaying instructions on specific zones, helping workers interact more intuitively with robotic arms and autonomous guided vehicles (AGVs).

2.3.4.3 Visual Feedback and Real-Time Interaction

One of the most significant advantages of AR-enhanced visual interaction is the provision of real-time feedback, which facilitates more intuitive robot control. [13] investigates the real-time testing of vision-based systems in AGVs using ArUco markers and AR visualizations. Their research highlights how visual feedback improves decision-making in dynamic environments, helping robots navigate autonomously with greater efficiency.

Visual feedback can take various forms:

- **Control Zones:** Specific areas within the robot's environment are highlighted using AR, guiding the robot's actions and helping the user visualize where the robot can move or interact with objects.
- **Status Updates:** Information such as battery levels, current tasks, or environmental hazards can be projected onto the robot's workspace, allowing users to quickly assess and adjust the robot's behavior as needed.
- **Error Detection:** AR systems can visually highlight potential errors in the robot's movements, such as collisions or off-course navigation, enabling quick corrective actions.

2.3.4.4 Benefits and Limitations

[8] emphasizes that visual interaction not only enhances HRI but also promotes collaboration between human operators and robots by reducing cognitive load. By providing real-time visual cues, AR improves situational awareness and makes controlling robots more intuitive, particularly in environments where precision is critical.

However, there are certain limitations. [12] discusses the challenges of accurately overlaying visual data in real-time, particularly in environments with inconsistent lighting, reflections, or

occlusions. Additionally, the processing demands of rendering real-time visualizations may limit the performance of systems with constrained computational resources.

2.4 AR’S ROLE IN IMPROVING USER ENGAGEMENT

Augmented Reality (AR) has significantly enhanced user engagement with robotic systems, particularly in Automated Guided Vehicles (AGVs). By providing intuitive visual feedback and real-time streaming data, AR improves both the psychological and practical interaction between humans and AGVs. The combination of augmented reality and robotics has been a topic of significant research interest, as AR makes robots more intuitive to control and increases user trust in automated systems.

2.4.1 Enhancing User Engagement through Real-Time Feedback and Visualization

One of the key strengths of AR is its ability to provide real-time feedback and visualizations, allowing users to intuitively understand the robot’s actions and status. AR-infused streaming technology improves human-robot collaboration by projecting real-time data, control zones, and navigation paths onto the robot’s environment [15]. This continuous stream of real-time data reduces cognitive load, enabling users to anticipate and control AGV movements more effectively in dynamic environments. Such real-time visualization allows users to engage more deeply with robotic systems, increasing their confidence in controlling AGVs.

In particular, [15] explore how AR is utilized across different robotic applications, including medical, industrial, and social robotics. For AGVs, they highlight AR’s capability to enhance human-robot collaboration by improving real-time task allocation and decision-making processes. The visual feedback offered by AR systems reduces the need for constant manual adjustments, thereby increasing operator efficiency [15].

In industrial scenarios, [16] emphasize the role of AR in improving user engagement through its intuitive interfaces. They note that AR helps operators visualize AGV tasks more clearly, thus reducing the learning curve for new users while providing advanced feedback for experienced operators. This is particularly important in fast-paced environments where quick decisions are necessary [16].

2.4.2 Psychological and Practical Benefits

The psychological impact of AR on user engagement is considerable. Real-time visualizations, such as path projections or task instructions, create a more predictable and transparent interaction between humans and robots. Users become more confident when they can clearly see an AGV's next move, which reduces uncertainty about the robot's behavior and enhances user trust. This **increased predictability** improves user satisfaction and promotes wider adoption of AR-enabled robotic systems [11].

AR also offers practical benefits in rehabilitation settings, where users can engage with robotic systems in therapy. As Makhataeva [16] points out, AR interfaces allow users to interact with AGVs or rehabilitation robots in a natural way, improving patient recovery outcomes. The immersive nature of AR encourages patient involvement, which is crucial in therapeutic contexts [16].

2.4.3 Applications in AGV Control

AR's integration into AGVs has numerous practical benefits, especially in logistics and warehousing. AR systems enable operators to visually define control zones for AGVs, making it easier to coordinate the movement of goods. For example, operators can dynamically adjust AGV routes using AR markers or floor projections, as explored by [15]. This capability increases flexibility and efficiency in fast-paced environments like warehouses.

Moreover, the AR-infused streaming technology enhances user engagement by displaying key metrics such as speed, battery levels, and task progress directly in the user's field of view. This eliminates the need to consult separate devices, improving operational efficiency and reducing the likelihood of errors in AGV management [10].

2.4.4 Conclusion

Augmented Reality has proven to be a powerful tool for improving user engagement with AGVs. By offering real-time feedback, intuitive visualizations, and immersive interactions, AR systems enhance both the psychological comfort and practical efficiency of users. This fosters greater trust in robotic systems, making them more approachable and easier to control. In environments such as manufacturing, healthcare, and logistics, AR-infused systems help users interact more naturally with AGVs, leading to improved task performance and satisfaction.

2.5 SENSOR TECHNOLOGIES IN HUMAN-ROBOT INTERACTION

When discussing the AR technology, it's difficult to do so without making mention of sensors, allowing machines to interact with their environments and humans. Sensors are to machines what eyes and ears are to humans. This section serves as a review for the applicable sensors relevant to this project which is designing a AR framework for a 4WD robotic car.

2.5.1 Types of Sensors for Mobile Robots

2.5.1.1 Ultrasonic Sensors

Ultrasonic sensors are a common choice in mobile robotics due to their effectiveness in obstacle avoidance and distance measurement. These sensors emit ultrasonic waves and measure the time it takes for the waves to reflect back from an object, allowing the robot to calculate the distance from obstacles [17]. This capability provides spatial awareness that is critical for safe navigation. Below is a basic diagram showcasing the operation of an HC-SR04 Infrared Sensor.

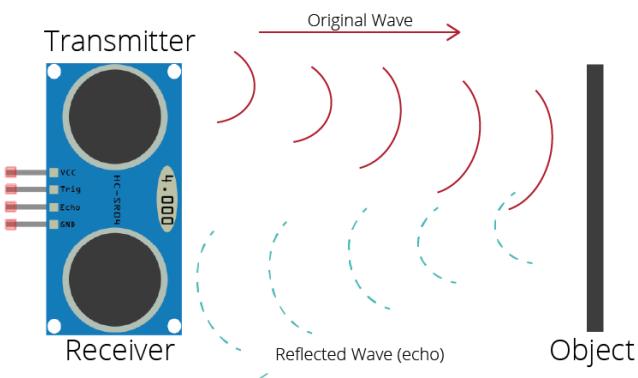


Figure 2.4: Operation of an Ultrasonic Sensor [1]

Below is a comparison of some common ultrasonic sensors used in mobile robotics.

Table 2.2: Comparison of Ultrasonic Sensors

Model	Range (cm)	Operating Voltage (V)	Accuracy	Field of View	Size (mm)
HC-SR04	2-400	5	± 0.3 cm	15 degrees	45 x 20 x 15
Maxbotix MB1000	20-645	2.5-5.5	$\pm 1\%$	42 degrees	22 x 20 x 16
Parallax PING)))	3-300	5	± 0.5 cm	20 degrees	22 x 46 x 16

Properties and Benefits for the Project

Ultrasonic sensors are advantageous because of their simplicity, affordability, and ease of integration. For this project, they provide an excellent mechanism for detecting nearby obstacles, which is essential for navigation in dynamic environments. Some key properties of ultrasonic sensors include:

- **Wide Range of Detection:** As seen in the table, ultrasonic sensors like the HC-SR04 and Maxbotix MB1000 can detect objects at distances ranging from a few centimeters to several meters. This is beneficial for both short-range and long-range obstacle avoidance.
- **Accuracy and Precision:** Depending on the model, ultrasonic sensors offer good accuracy for distance measurements. For example, the Maxbotix MB1000 boasts an accuracy of $\pm 1\%$, helping ensure that the robot reacts promptly and effectively to its surroundings.
- **Field of View (FoV):** The field of view differs between sensors, affecting the robot's perception. A larger FoV allows the robot to detect obstacles from a wider angle, but may result in less precise focus on individual objects.
- **Interfacing and Visualization:** Integrating ultrasonic sensors with augmented reality (AR) offers the possibility of visualizing sensor data in real time. This can provide users with clear, dynamic feedback on the robot's perception of obstacles, displayed as visual cues through AR devices. By overlaying distance and obstacle data on the robot's surroundings, the user can have an immersive understanding of the environment, improving control and decision-making during operation.

For example, AR can highlight areas where the robot detects obstacles, providing the operator with enhanced spatial awareness, preventing collisions, and improving task efficiency. Additionally, real-time data streaming through a visual interface allows for immediate adjustments to the robot's behavior based on sensor readings.

2.5.1.2 Infrared Sensors

Infrared (IR) sensors are commonly used in mobile robotics for tasks such as proximity detection, line-following, and object avoidance. These sensors emit infrared light and detect the reflection from nearby objects, making them effective for short-range detection in controlled environments.

While standard IR sensors, such as the basic proximity IR sensors, serve well in obstacle avoidance and object detection, they have limitations in terms of range, field of view, and accuracy.

These standard sensors perform well in specific lighting conditions but struggle in environments with significant ambient light interference.

2.5.1.2.1 The Dagu Infrared Compound Eye: An Enhanced Solution The Dagu Infrared Compound Eye shares basic functionality with standard IR sensors but offers significant enhancements. Unlike single-point IR sensors, the Dagu sensor features a compound array of multiple infrared detectors, allowing for a much wider field of view and multi-directional sensing. This enables the robot to detect objects from multiple angles, enhancing spatial awareness and enabling more complex navigation and object tracking tasks.

2.5.1.2.2 Advantages for This Project Integrating the Dagu Infrared Compound Eye into this project would extend the robot's capabilities by improving its obstacle detection, particularly in dynamic and complex environments. The sensor's wide field of view, combined with real-time AR feedback, could provide users with a more intuitive control system and greater situational awareness, thereby improving human-robot interaction (HRI). Moreover, the sensor could aid in tasks requiring precise spatial awareness, such as navigating around obstacles or detecting specific objects.

Table 2.3: Specifications of the Dagu Infrared Compound Eye

Specification	Value
Number of Detectors	4 infrared sensors
Detection Range	200mm
Operating Voltage	5V
Power Consumption	10 mA

Additionally, this sensor could provide valuable data that could be visualized for the user through augmented reality (AR), enabling a more interactive and user-friendly experience.

2.5.1.3 LiDAR Sensors

LiDAR (Light Detection and Ranging) is an essential technology for mobile robotics that uses laser light to measure distances and create 3D maps of environments. By utilizing LiDAR sensors, mobile robots can detect obstacles, map their surroundings, and navigate autonomously.

2.5.1.3.1 RPLIDAR A1 and TF Mini LiDAR Sensors For this project, two LiDAR sensors will be considered: **RPLIDAR A1** and **TF Mini LiDAR (ToF) Laser Range Sensor V2.0**. These sensors offer complementary capabilities that can significantly enhance the

robot's ability to perceive and interact with its environment. Below is a comparison of their specifications:

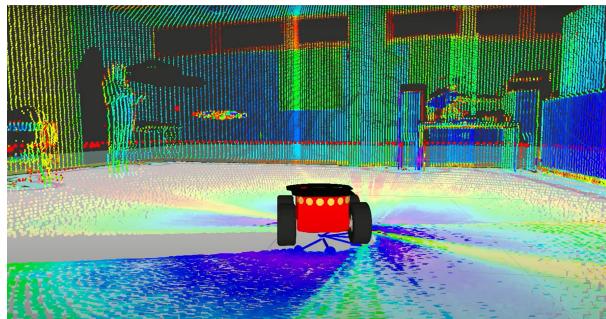


Figure 2.5: LiDAR scanning and mapping process in mobile robotics.

Table 2.4: Comparison of RPLIDAR A1 and TF Mini LiDAR (ToF) Sensors

Specification	RPLIDAR A1	TF Mini LiDAR (ToF)
Max Range	12m	12m
Resolution	0.2cm	1cm
Field of View	360° (horizontal)	3.6° (horizontal)
Sample Rate	8000 samples/sec	1000 samples/sec
Power Consumption	5V, 500mA	5V, 140mA
Interface	UART, USB	UART
Size (mm)	97.5 x 60 x 38	42 x 15 x 16

The **RPLIDAR A1** is a 360-degree scanning LiDAR sensor capable of creating a comprehensive map of the robot's surroundings in real time. This would enable the robot to detect objects and obstacles from all directions, making it ideal for use in complex environments where autonomous navigation is crucial.

On the other hand, the **TF Mini LiDAR (ToF)** sensor, although more limited in its field of view, offers a compact and lightweight design, making it suitable for applications requiring focused, high-speed distance measurement. Its small size and low power consumption could be beneficial for tasks where precise distance measurements are needed without adding significant weight or power demands.

Both sensors would improve the robot's environmental mapping and obstacle detection capabilities. Visualizing the LiDAR data in an AR interface would provide real-time feedback to users, enabling them to monitor the robot's surroundings and make informed decisions about its movement.

2.6 Web-Based Control Interfaces for Mobile Robots This section will detail different methods of implementing web-based control interfaces, focusing on real-time communication, user interaction, and video streaming.

2.6.1 Overview of Web Server Technologies Explain the importance of web-based interfaces for human-robot control and how various frameworks make it easier to send commands and stream video.

2.6.2 Webserver Framework Options Flask: Discuss its lightweight nature and ease of use for rapid web-based control interface development, especially for real-time command streaming. Node.js: Highlight its asynchronous, non-blocking I/O operations, making it suitable for real-time control interfaces that demand quick responses. WebSocket Integration: Discuss the need for low-latency, full-duplex communication between the server and the robot. Web-Sockets could be used for sending control commands in real-time and receiving video streams. Apache/Nginx: Compare these more traditional web servers, emphasizing their scalability and robustness, but also mention their complexity in smaller, embedded systems like a Raspberry Pi. 2.6.3 Security and User Experience Discuss how real-time control interfaces can improve HRI, but also introduce the challenges of ensuring secure communication between the robot and the user. You could also expand on how well-designed web interfaces enhance user experience, particularly when combined with AR.

Integrating these topics into the broader literature review To integrate these sections seamlessly into your existing structure, consider how each relates to the previous discussions:

Sensor technologies tie into fiducial marker systems in AR by providing data about the robot's surroundings. Both improve task execution, interaction, and navigation in dynamic environments, which aligns with the project's goals of enhancing user experience through AR.

Web server setups complement the section on AR's role in user engagement by discussing the technical details of how control is transmitted from the human to the robot. A well-implemented web interface ensures the smooth operation of the HRI system and supports the project's goal of intuitive, real-time control.

This will enhance your literature review, showing how technology components (sensors, web servers) work together in an AR-enabled HRI system.