ECE 445

SENIOR DESIGN LABORATORY

PROPOSAL

Robotic Arm Integrated into Wheelchair with MR Interface

Team #12

XINGRU LU (xingrul2@illinois.edu)
YILIN WANG (yilin14@illinois.edu)
YUNYI LIN (yunyil3@illinois.edu)
YINUO YANG (yinuoy4@illinois.edu)

TA: Yun Long

March 14, 2025

Contents

1 Introduction **1**
1.1 Objective and Background . . . 1
1.1.1 Goals . . . 1
1.1.2 Functions . . . 1
1.1.3 Benefits . . . 1
1.1.4 Features . . . 1
1.2 High-Level Requirements List . . . 2

2 Design and Requirements **3**
2.1 Diagrams . . . 3
2.2 Robotic Arm Module . . . 5
2.2.1 Power . . . 5
2.2.2 Customized End Effector . . . 5
2.2.3 Open Manipulator-P Robotic Arm . . . 5
2.3 Mixed Reality Module . . . 6
2.3.1 Depth Camera . . . 6
2.3.2 Mixed Reality Interface System . . . 6
2.3.3 Apple Vision Pro . . . 7
2.4 Mobile Platform Module . . . 7
2.5 Tolerance Analysis . . . 7
2.5.1 Restricted Raw Vision Pro Data Access . . . 7
2.5.2 Compatibility Issues with ROS1 and External Depth Cameras . . . 8
2.5.3 Head Movement Impact on Hand Tracking Accuracy . . . 8
2.5.4 Latency in the System Pipeline . . . 8

3 Ethics and Safety **10**
3.1 Robotic Arm Operation Safety . . . 10
3.2 Privacy and Data Protection . . . 10
3.3 User Autonomy and Accessibility . . . 10

References **11**

1 Introduction

1.1 Objective and Background

1.1.1 Goals

Wheelchair users often face significant challenges when interacting with objects beyond their immediate reach, particularly behind them. Without external assistance, tasks such as pressing buttons or navigating through environments with complicated surroundings can become difficult. These difficulties are compounded when operating independently, highlighting the need for supplementary support to simplify routine activities. Additionally, wheelchair users may struggle with limited situational awareness, as their field of view is primarily forward-facing. As a result, there is a pressing need for innovative solutions that enhance both accessibility and autonomy, enabling wheelchair users to interact more conveniently with their surroundings.

1.1.2 Functions

Our solution integrates a rear-facing camera that streams real-time visuals to a Mixed Reality (MR) interface, allowing wheelchair users to gain visual awareness of their surroundings, including blind spots behind them. Additionally, a robotic arm mounted at the back of the wheelchair can be controlled through MR, enabling users to perform assistive actions such as pressing buttons and interacting with objects beyond their physical reach. This system enhances both situational awareness and independent mobility, providing a more intuitive and convenient way for users to navigate and interact with their environment.

1.1.3 Benefits

Our solution helps people with wheelchairs by enhancing **situational awareness** and **independent mobility**. It provides real-time imagery with the Mixed Reality (MR) interface via the rearview camera, allowing users to see blind spots and navigate more securely. Additionally, Users can use the robotic arm, controlled by MR, to push buttons, for example, to make it easier to perform everyday tasks independently. It gives users an easier and more intuitive way to interact with the environment, ultimately resulting in greater autonomy and accessibility.

1.1.4 Features

- **Mixed Reality Integration:** Combines real-time MR with assistive robotics, offering an intuitive control experience.
- **Enhanced Situational Awareness:** A rear-facing camera provides live visuals of blind spots, improving navigation.
- **Extended Reach with Robotic Arm:** The Open Manipulator-P allows users to interact with objects beyond their grasp, such as pressing buttons behind them.

- **Apple Vision Pro for Control & Feedback:** Tracks hand movements for precise robotic control and provides interactive feedback.
- **Low Latency & Safety Focus:** Ensures smooth operation without compromising wheelchair stability.

1.2 High-Level Requirements List

- **Precision:** The robotic arm should reliably press buttons with a diameter of at least **35mm**, which is a common size of elevator buttons. The force applied must be sufficient to activate buttons without excessive pressure that could cause damage or failure.
- **Safety and Stability:** Users should be able to see both the front and rear environments through Vision Pro, while also adjusting the robotic arm's perspective to gain a broader field of view.
- **Reach:** The robotic arm should be able to reach a height from 110cm to 160cm.

2 Design and Requirements

2.1 Diagrams

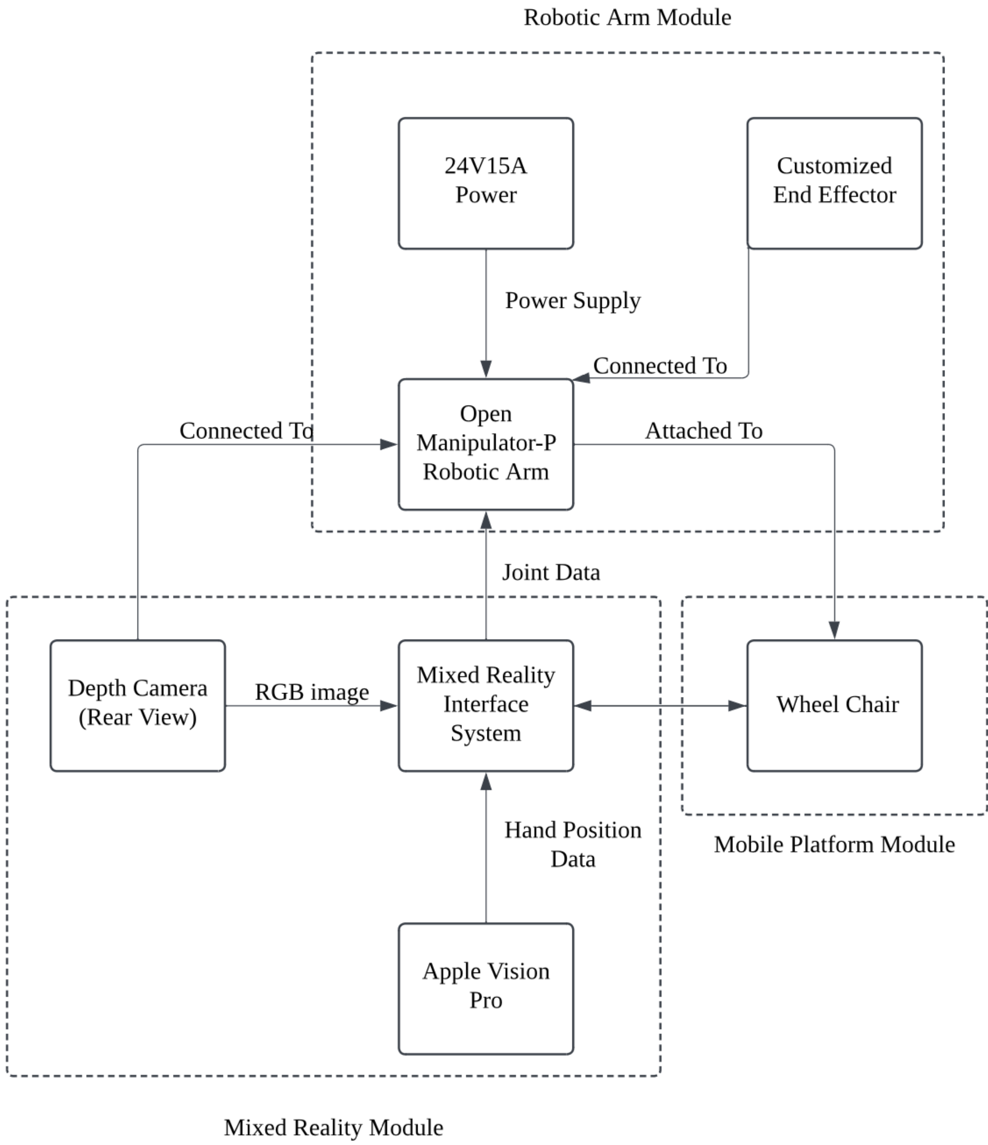

Figure 1: Block Diagram

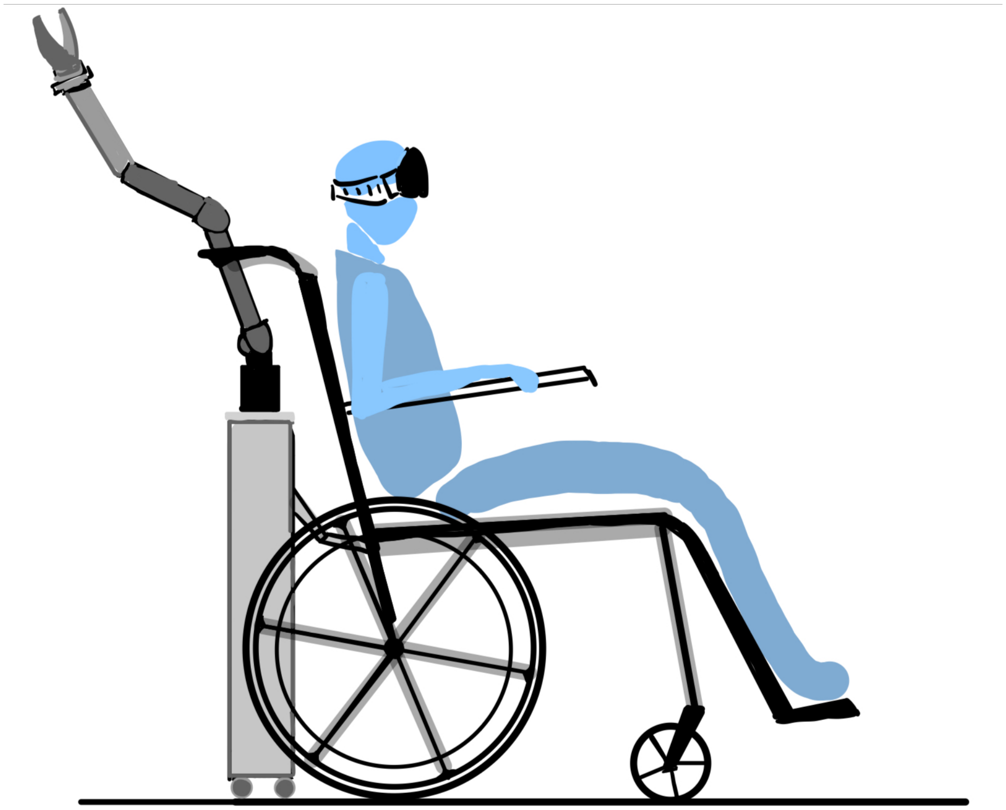

Figure 2: Physical Overview

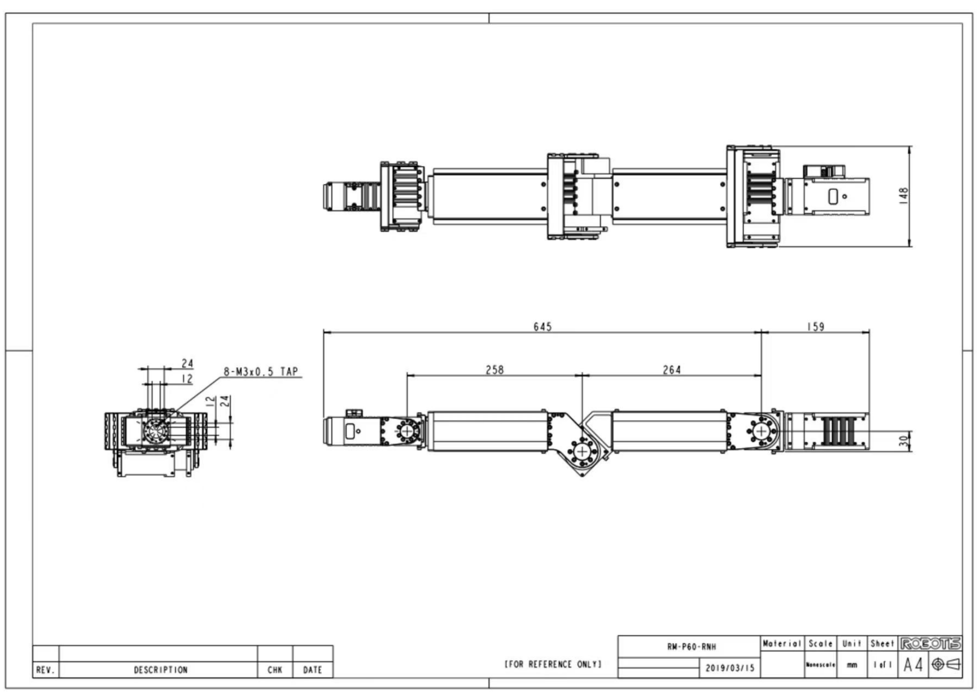

Figure 3: Dimensions of OpenManipulator-P [1]

2.2 Robotic Arm Module

The Robotic Arm Module consists of an Open Manipulator-P Robotic Arm, powered by a 24V15A Power Supply and equipped with a Customized End Effector to interact with the environment, such as to press buttons. It is physically mounted on the Mobile Platform Module (Wheelchair) and connected to the Mixed Reality Module for feedback and control.

2.2.1 Power

The Robotic Arm Module consists of an Open Manipulator-P Robotic Arm, powered by a 24V15A Power Supply and equipped with a Customized End Effector to interact with the environment, such as to press buttons. It is physically mounted on the Mobile Platform Module (Wheelchair) and connected to the Mixed Reality Module for feedback and control.

2.2.2 Customized End Effector

The Customized End Effector is a gripper-style clamp designed for basic interaction tasks, such as pressing buttons, supporting tasks that require moderate accuracy but not extreme precision. It is mounted mechanically on the Open Manipulator-P Robotic Arm and operates based on the signals provided by the Mixed Reality Interface System.

Requirements:

- The end effector is optimized for pressing buttons with a minimum diameter of 35mm.

2.2.3 Open Manipulator-P Robotic Arm

The Open Manipulator-P Robotic Arm is a 6-DOF robotic arm with a 645mm reach, is the principal actuator for the system. It executes commands from the Mixed Reality Interface System to interact with the environment and is physically attached to the Mobile Platform Module (Wheelchair). Powered by a 24V15A Power Supply, the arm operates in conjunction with the Customized End Effector to perform physical tasks. It is designed for moderate-speed performance with controlled, smooth motion to avoid instability.

Requirements:

- The robotic arm should compensate for cumulative joint errors to keep the end-effector error within the 5mm threshold.
- The force applied to the button should remain within the required activation force range (typically 1-5N, depending on the button type) to prevent missed presses or excessive force.

2.3 Mixed Reality Module

The Mixed Reality Module consists of the Apple Vision Pro, the Depth Camera, and the Mixed Reality Interface System, the core processor for interpreting inputs provided by the user and information regarding the surroundings. It sends precise commands to the Robotic Arm Module via joint data, enabling intuitive and adaptive control. The Depth Camera expands the system's field of view, and the Mixed Reality Interface System maps the motion of the user's hand for intuitive interaction with the robotic arm.

2.3.1 Depth Camera

The Depth Camera (Rear View) captures RGB images and depth information, providing extra environmental awareness for the user. It's a single depth camera mounted at the rear of the system. This data supports accurate robotic arm movements and improves user control over the whole system.

Requirements:

- The Depth Camera (Rear View) must have a minimum resolution of 1280×720 pixels to ensure clear object recognition.
- The camera's depth accuracy must be within ± 10mm to provide reliable distance estimation.
- The data processing delay should not exceed 200ms to maintain real-time feedback for accurate system responses.
- The camera must have a horizontal Field of View (FoV) of at least 90° to cover a sufficient environmental area.

2.3.2 Mixed Reality Interface System

The Mixed Reality Module consists of the Apple Vision Pro, the Depth Camera, and the Mixed Reality Interface System, the core processor for interpreting inputs provided by the user and information regarding the surroundings. It processes hand tracking data, depth information, and environmental cues to enable intuitive, real-time control, allowing users to interact seamlessly with the system through mixed reality gestures.

Requirements:

- The system must process user input within 150ms to ensure a responsive and natural interaction experience.
- The system must maintain data synchronization latency under 50ms between sensor inputs and control outputs to prevent motion lag.
- It should handle at least 30 frames per second (FPS) for smooth and continuous mixed reality interaction.

2.3.3 Apple Vision Pro

The Apple Vision Pro serves as the wearable input device for real-time hand tracking and gesture-based control in the Mixed Reality Module. It captures precise hand position data and transmits it to the Mixed Reality Interface System, enabling intuitive user interaction with the Robotic Arm Module and Mobile Platform Module (Wheelchair). By making use of spatial computing and mixed reality capabilities, the Apple Vision Pro allows users to control the system with natural hand movements, eliminating the need for physical controllers.

Requirements:

- The hand tracking accuracy should be within ± 5mm to ensure precise robotic control.
- The gesture recognition latency should not exceed 100ms, preventing delays between user input and system response.
- The tracking refresh rate must be at least 60 FPS to maintain smooth and continuous interaction.
- The field of view (FoV) for hand tracking should cover at least 90° horizontally to allow comfortable and unrestricted gestures.
- If tracking deviation exceeds 10mm or latency surpasses 150ms, the system should trigger recalibration or notify the user of tracking issues.

2.4 Mobile Platform Module

The Mobile Platform Module is the foundation for the mobility in the system, allowing users to navigate the space while maintaining control of the Robotic Arm Module. It's mainly a wheelchair, which is controlled by the Mixed Reality Interface System to make the navigation either speech-based or gesture-based. When integrated with the Robotic Arm Module, the platform allows users to work with objects even when moving, enhancing independence and accessibility.

2.5 Tolerance Analysis

2.5.1 Restricted Raw Vision Pro Data Access

One of the most significant hurdles to the application of hand tracking with Vision Pro is the limits on compatibility and permission that Apple's environment imposes. While Vision Pro enables the possibility for high-accuracy hand tracking by utilizing ARKit, third-party applications are generally excluded from direct access to raw camera data by Apple's security and privacy policies. This limit could constrain the level of detail available for external processing. Apple's strict security policies may also prevent direct access to low-level hand tracking data, requiring developers to work within the bounds of available APIs.

2.5.2 Compatibility Issues with ROS1 and External Depth Cameras

While Vision Pro provides advanced hand-tracking functionality, our robotic arm can rely on an onboard depth camera for control and sensing. Many modern depth cameras and vision-based tracking architectures, however, are optimized for use with ROS2, whereas our robotic arm is best supported by ROS1. This incompatibility can result in issues with the communication of the data, driver support, and real-time processing between the depth camera and the robotic control system based on ROS1.

2.5.3 Head Movement Impact on Hand Tracking Accuracy

When the hand is being tracked with Vision Pro, the motion of the head-mounted device (HMD) can make the position of the hand appear to be incorrect. Since the tracking is relative to the position of the hands with respect to the headset, motion of the headset alters the reference frame and can lead to differences between the actual and perceived hand positions.

Let:

- $\mathbf{H}$ be the headset coordinate frame,
- $\mathbf{W}$ be the world coordinate frame,
- $\mathbf{P}_{h|H}$ be the hand coordinate frame (relative to the headset),
- $\mathbf{T}_{H|W}$ be the headset movement relative to the world coordinate system,
- $\mathbf{T}_{h|H}$ be the hand movement relative to the headset,
- $\mathbf{T}_{h|W}$ be the actual hand movement in the world coordinate system.

The actual hand displacement in the world frame should be calculated as:

$$\mathbf{T}_{h|W} = \mathbf{T}_{H|W} + \mathbf{T}_{h|H}$$

However, if the system does not compensate for head movement, it assumes:

$$\mathbf{T}_{h|W} = \mathbf{T}_{h|H}$$

which leads to errors in hand position estimation. This discrepancy is especially problematic when the user performs fine motor control tasks, where precise tracking is critical.

2.5.4 Latency in the System Pipeline

Latency is a critical factor in the real-time control of a robotic arm using Vision Pro hand tracking. The total system latency consists of four main components:

- **Vision Pro Hand Tracking Latency:** The time delay between the user's actual hand movement and the system's recognition of the new hand position. According to Road to VR, Vision Pro's measured hand tracking latency is approximately 128 ms **robertson2024**.
- **Program Processing Latency:** The time required for the computing system to process the received hand position data and generate corresponding control commands for the robotic arm. This depends on computational complexity, data transmission speed, and software optimizations.
- **USB Communication Latency (DYNAMIXEL U2D2):** As noted in the DYNAMIXEL U2D2 e-Manual, when connecting the robotic arm to the PC via USB, the default USB latency time is 16 ms.
- **Robotic Arm Reaction Latency:** The delay between receiving the control command and the robotic arm physically executing the movement. This includes actuation delay, motor inertia, and mechanical constraints.

Thus, the total system latency can be expressed as:

$$T_{\text{total}} = T_{\text{hand tracking}} + T_{\text{processing}} + T_{\text{USB}} + T_{\text{robot arm}}$$

Substituting known values:

$$T_{\text{total}} = 128 + T_{\text{processing}} + 16 + T_{\text{robot arm}}$$

For real-time interaction, it is essential that T_{total} remains below the human perceptual threshold for smooth interactions, typically around 200-250 ms for motion feedback applications. If T_{total} exceeds this limit, users may experience noticeable lag, affecting the precision and responsiveness of robotic control.

If we assume:

$$T_{\text{processing}} \approx 50 \text{ ms}, \quad T_{\text{robot arm}} \approx 40 \text{ ms}$$

Then the total system latency falls within:

$$T_{\text{total}} = 128 + 50 + 16 + 40 = 234 \text{ ms}$$

which is within the acceptable range.

3 Ethics and Safety

3.1 Robotic Arm Operation Safety

Our system includes a robotic arm which extends from the rear of the wheelchair, which introduces potential risks of harming the surrounding public if not properly designed and tested. To uphold the IEEE principle of prioritizing public safety and well-being, we will implement hardware and software safeguards. When not in the process of reaching, the robotic arm will remain in a folded position and not occupy more horizontal space than the wheelchair itself. While in motion, the speed of the robotic arm will be constrained so as to not pose risk of heavy collision with objects. The Open Manipulator-P robotic arm will be programmed to ensure that it operates within safe limits and does not pose a risk to the user or nearby individuals. Furthermore, Apple Vision Pro's depth and spatial awareness capabilities will be utilized to enhance situational awareness and prevent unintended interactions.

3.2 Privacy and Data Protection

User privacy is a critical consideration in our system, particularly given the use of real-time cameras and Mixed Reality (MR) technology. Our system does not store or transmit user data to any external servers. All video processing and interaction tracking occur locally. The rear-facing camera feed is processed in real time solely for user awareness and robotic arm control. Similarly, Apple Vision Pro's hand-tracking data is processed locally, without transmitting biometric or movement data beyond the device [2].

3.3 User Autonomy and Accessibility

Our solution is designed to empower wheelchair users by enhancing their autonomy, which aligns with the IEEE and ACM principles of respecting human dignity and promoting inclusivity [3], [4]. The MR interface will be developed with intuitive controls, allowing users to interact comfortably with the system. While providing rear-view video feedback to user, the MR interface should not obscure the frontal view of the user and ensure safety.

By integrating these ethical and safety considerations, our project aims to enhance the autonomy of wheelchair users while ensuring their safety and privacy.

References

[1] ROBOTIS. “Openmanipulator-p - specification,” ROBOTIS e-Manual. (Accessed: Mar. 14, 2025), [Online]. Available: https://emanual.robotis.com/docs/en/platform/openmanipulator_p/specification/#specifications (visited on 03/14/2025).

[2] Apple Inc. “Apple vision pro privacy overview,” Apple. (2023), [Online]. Available: https://www.apple.com/privacy/ (visited on 03/14/2025).

[3] IEEE. “Ieee code of ethics,” Institute of Electrical and Electronics Engineers. (2020), [Online]. Available: https://www.ieee.org/about/corporate/governance/p7-8.html (visited on 03/14/2025).

[4] ACM. “Acm code of ethics and professional conduct,” Association for Computing Machinery. (2018), [Online]. Available: https://www.acm.org/code-of-ethics (visited on 03/14/2025).