Smart Voice-Based AI Robot

Under the guidance of

Dr. Kaveri

Professor & Head

Department of Electronics and Communication Engineering

Aditya College of Engineering and Technology

Bangalore, India

Ashwini H Gummadi

Department of Electronics and Communication Engineering

Aditya College of Engineering and Technology

Bangalore, India

ashwinigummadi123@gmail.com

Hamsa Shree

Department of Electronics and Communication Engineering

Aditya College of Engineering and Technology

Bangalore, India

grhamsa2005@gmail.com

Ekshitha R

Department of Electronics and Communication Engineering

Aditya College of Engineering and Technology

Bangalore, India

ekshithar297@gmail.com

Himavardhan L

Department of Electronics and Communication Engineering

Aditya College of Engineering and Technology

Bangalore, India

saihimavardhan001@gmail.com

ABSTRACT

The project titled **Smart Voice-Based AI Robot** aims to design and implement an intelligent robotic system that operates according to voice instructions provided by the user. The main objective is to develop a robot that can interact with humans through voice commands, reducing the need for traditional manual operation. The ESP32-S3 development board acts as the central processing unit and coordinates the activities of different hardware components. A microphone is used to receive voice inputs, which are interpreted by the control system to execute movements such as forward, backward, left, and right.

For controlling the robot's movement, a motor driver is interfaced with the controller. A servo motor is included to support directional movement, while an ultrasonic sensor continuously checks the surrounding area for obstacles. A speaker is also integrated to provide audio feedback, making communication between the user and the robotic system more interactive. These components work together to achieve voice-operated movement and basic obstacle detection.

This project demonstrates the application of embedded systems and intelligent control techniques in robotics. The developed model can serve as a learning platform for automation and human-robot interaction. In the future, its capabilities can be expanded by incorporating improved speech-processing methods, machine learning algorithms, natural language understanding, and other intelligent features.

Keywords: Voice-Operated Robotics, ESP32-S3 Controller, Speech Recognition, Intelligent Automation, Obstacle Avoidance, Ultrasonic Sensing, Embedded Technology, Human-Robot Interaction.

I. INTRODUCTION

The advancement of robotics and intelligent computing has created new opportunities for developing automated systems that can interact with humans. Modern robotic applications are increasingly designed to perform operations with minimal manual involvement. Among the available interaction methods, voice communication offers a convenient way to operate machines because users can provide instructions through speech rather than depending on physical switches or handheld controllers.

A voice-operated robotic system works by receiving spoken commands and converting them into instructions that the controller can understand. Speech recognition plays an important role in this process by identifying the words spoken by the user and associating them with specific actions. Using this approach, a mobile robot can be directed to move forward, move backward, turn left, or turn right. Several research studies have investigated speech-controlled robotic platforms and their ability to improve interaction between humans and machines. Nevertheless, recognition performance can be influenced by environmental noise, variations in pronunciation, and the accuracy of the speech-processing method.

For a mobile robot to operate effectively, movement control must be supported by appropriate sensing mechanisms. Obstacle detection is an important feature that helps identify objects in the robot's surroundings. An ultrasonic sensor performs this function by measuring the distance between the robot and nearby objects. The information obtained from the sensor can be used by the controller to determine an appropriate response when an obstacle is detected. Earlier developments in voice-controlled robotics have also considered the combination of speech commands and ultrasonic sensing for improving operational safety. Microcontrollers make this integration possible by managing communication between the sensing units, movement mechanisms, and other electronic devices.

The present project introduces a **Smart Voice-Based AI Robot developed around the ESP32-S3 microcontroller**. The microcontroller acts as the central unit responsible for processing inputs and coordinating the connected hardware. Voice instructions are received through a microphone and processed by the programmed control system. The motor driver supplies the necessary control signals for robot movement, while a servo motor is used for directional adjustment. An ultrasonic sensor monitors the surrounding area for obstacles, and a speaker provides audio feedback during operation. These components are integrated to achieve a robotic platform capable of responding to basic directional commands.

Artificial Intelligence has expanded the capabilities of electronic systems by introducing methods for processing information and interpreting user requirements. Technologies such as speech recognition and Natural Language Processing help establish communication between humans and intelligent devices. Their integration into robotic applications makes it possible to develop systems that respond to spoken instructions in a more accessible manner.

The proposed robotic system focuses on connecting voice-based instructions with corresponding physical operations. Depending on its software implementation, the platform can also be extended to support additional functions such as answering basic queries, providing spoken information, and operating other connected devices. The speaker provides an additional

communication facility by allowing the system to deliver audio responses to the user.

The primary aim of this work is to demonstrate how embedded controllers, voice interaction, and obstacle sensing can be combined to develop an interactive robotic application. The system provides a practical learning platform for understanding intelligent automation and human-machine communication. Its design can be further adapted for educational activities, assistance-related applications, and simple automated environments. Future improvements may include more accurate speech processing, advanced decision-making techniques, and additional intelligent features.

II. LITERATURE SURVEY

The use of voice commands in robotics has become an important area of research due to improvements in speech-processing technologies and embedded computing. Researchers have investigated several approaches to make robotic systems respond to human instructions more effectively. These approaches include speech recognition, wireless control, intelligent navigation, and sensor-based obstacle identification.

A. Mobile Robot Operation Through Voice Commands (2001)

T. Kubik and M. Sugisaka studied the possibility of operating mobile robotic systems through natural speech instructions. Their research focused on the use of mobile communication devices and voice-processing techniques for transmitting commands to robots. The study indicated that the quality of incoming audio information is an important factor in obtaining accurate control responses. This work contributes to the early development of speech-operated robotic applications.

B. Speech-Based Robot Navigation with ROS (2024)

M. A. Kasno et al. investigated a voice-assisted navigation technique implemented on the TurtleBot3 robotic platform. The system was developed using ROS Noetic, where spoken instructions were collected through a microphone and interpreted by speech recognition software. After identifying the requested operation, the system generated suitable movement commands for the robot, including forward, reverse, left, right, and stop. The researchers also examined the influence of surrounding noise and the importance of improving speech identification. Their work demonstrates the application of robotic software frameworks in voice-based navigation.

C. Integration of Voice Commands and Obstacle Detection (2025)

B. I. Reddy et al. presented a mobile robotic system that combines speech-based operation with distance-sensing technology. Their implementation used an Arduino-based controller along with a mobile application and Bluetooth communication for receiving user instructions. An ultrasonic sensor was incorporated to monitor the area surrounding the robot and identify objects that could interfere with its movement. This approach demonstrates the usefulness of combining wireless voice control and environmental sensing in robotic applications.

D. Robot Navigation Through Offline Voice Processing (2025)

W. B. Abdelaziz et al. examined the use of locally processed speech commands for controlling robotic navigation through the ROS2 framework. Their proposed arrangement converted spoken inputs into navigation instructions using an offline speech recognition mechanism. The performance of the approach was investigated in a simulated indoor environment. The study showed the relevance of offline speech processing in situations where continuous internet access is unavailable or unreliable.

E. Overall Findings from the Literature

The examined research works present different techniques for developing robots that can be operated through spoken instructions. Some studies concentrate on mobile communication and wireless command transmission, whereas others investigate ROS-based navigation, local speech processing, and sensor-assisted obstacle identification. These developments establish a technical foundation for

creating more accessible and interactive robotic platforms.

Although individual techniques have been explored in earlier studies, integrating voice-operated movement, obstacle monitoring, and sound-based feedback into a single embedded platform remains a useful project objective. The present work addresses this objective through an ESP32-S3-based robotic system. It combines a microphone, motor driver, servo motor, ultrasonic sensor, and speaker to support voice-controlled movement, obstacle detection, and communication with the user. This integration provides a foundation for developing a compact robotic assistant with scope for additional intelligent features.

III. METHODOLOGY

The development of the Smart Voice-Based AI Robot involves designing an electronic control system that connects voice instructions with robotic actions. The implementation is carried out through hardware assembly, controller programming, audio-input handling, command identification, and movement execution. The ESP32-S3 board functions as the main processing device, coordinating the connected modules and controlling the activities of the robot.

A. Hardware Configuration and Module Interfacing

The robotic prototype consists of an ESP32-S3 controller interfaced with several electronic components. The INMP441 microphone receives sound signals from the surrounding environment and supplies audio data to the controller. A motor driver is connected to the DC motors to regulate their operation and control the direction of movement. The servo motor is used to adjust the required position, whereas the ultrasonic sensor observes the area in front of the robot to identify nearby objects. An audio output unit is also provided to communicate responses to the user. All these modules are connected to form a coordinated robotic arrangement.

B. Audio Input and Speech Recognition Process

The interaction process starts when a user speaks a command near the microphone. The INMP441 sensor converts the received sound into digital audio information, which is supplied to the processing system. The speech recognition mechanism examines this information and determines the instruction given by the user. The identified command is then forwarded to the programmed controller, where it is associated with the corresponding robot operation. This process establishes voice communication between the user and the robotic platform.

C. Directional Movement Operation

Once the controller receives the identified instruction, it selects the required motor-control operation. The ESP32-S3 communicates with the motor driver through programmed output signals. These signals determine how the connected motors operate to produce the requested movement. The robot can perform basic navigation operations, including forward travel, reverse travel, and directional turns. The motor driver handles the electrical control required for operating the motors according to the selected command.

D. Recognition and Selection of Commands

The recognized speech information is checked against the predefined command set available in the program. The controller searches for the relevant instruction and selects the corresponding operation. Different commands are mapped to separate control functions so that the robot can distinguish between movement requests. This command-matching stage helps prevent unrelated input from being treated as a valid movement instruction.

E. Experimental Testing and Verification

After completing the hardware connections and software implementation, the prototype is subjected to a series of functional tests. Each movement instruction is supplied separately to examine the response of the robotic system. The operation of the motors is checked to confirm the required direction of movement. The ultrasonic sensor is evaluated using objects positioned at different distances to observe its ability to identify

nearby obstacles. The microphone, speaker, and controller are also examined to verify their individual functions. The results of these tests are used to assess the overall operation and identify areas that may require modification.

F. Wireless Interface and Connectivity

The robotic platform can support wireless interaction through communication technologies such as Wi-Fi and Bluetooth. These technologies provide a means of connecting the robot with compatible external devices. Depending on the communication method implemented, users can transmit control instructions without using a direct wired connection. This feature also creates opportunities to extend the robot with smartphone-based control and additional remote-operation facilities.

G. Overall Methodology Description

The proposed methodology establishes a connection between voice-based input and physical robotic operation. Audio information is collected through the microphone, interpreted by the recognition system, and converted into predefined instructions. The ESP32-S3 processes these instructions and coordinates the motor driver to produce the required movement. At the same time, the ultrasonic sensor provides information about nearby obstacles, and the speaker supports audio communication. This arrangement demonstrates the practical integration of embedded control, speech-based interaction, and obstacle monitoring in a single robotic prototype.

IV. PROPOSED SYSTEM

The Smart Voice-Based AI Robot is proposed as a voice-operated mobile robotic system that enables users to control its movements through spoken commands. The system reduces the requirement for conventional manual control methods, such as switches and remote-control devices. A mobile application is used to receive the user's voice input and identify the requested operation. The recognized instruction is converted into a predefined command and communicated to the robot through a wireless Bluetooth connection.

The ESP32-S3 microcontroller acts as the central control unit of the robotic vehicle. It receives the transmitted instructions, processes them, and activates the appropriate control operation. Depending on the command received, the motor driver regulates the DC motors to move the robot in the required direction. The system also supports audio communication through a speaker, which can deliver responses generated by the programmed application or controller. Text-to-Speech (TTS) functionality can be used wherever speech-based feedback is implemented.

This system integrates mobile application control, wireless communication, and embedded electronic components to achieve convenient robotic operation. By using voice instructions as the primary input method, it provides an alternative to traditional manual control. The proposed design can be used for academic demonstrations and as a starting point for developing assistive and automated robotic applications.

A. Main Features of the Proposed System

1. Voice Command Control: The system accepts spoken instructions through a mobile application and uses the recognized commands to control the robot's movement.

2. Intelligent Command Processing: The control program identifies the received instruction and selects the appropriate action from the predefined command set.

3. Speech-to-Text Recognition: The user's speech is converted into a readable text format or corresponding digital instruction for further processing.

4. Natural Language Understanding: NLP techniques can be added to improve the system's ability to interpret different forms of user instructions.

5. Automatic Movement Control: The robot carries out the selected movement by sending suitable signals from the controller to the motor driver.

6. Audio Response Facility: A speaker can be used to provide sound-based feedback. When TTS is

available, text responses can be converted into spoken output.

7. Wireless Communication: Bluetooth connectivity enables the mobile application to exchange control information with the ESP32-S3 without a physical data connection.

8. Additional Feature Support: The system architecture allows further development through functions such as advanced obstacle avoidance, object identification, and home automation.

B. Operational Description

The working process begins when the user provides a voice instruction through the mobile application. The application processes the spoken input and converts it into a corresponding predefined command. The command is then transmitted to the ESP32-S3 using Bluetooth Low Energy (BLE).

The microcontroller receives the wireless data and determines the required movement based on the programmed instructions. It sends appropriate signals to the motor driver, which controls the DC motors. Through this arrangement, the robot can move forward, reverse, turn left, or turn right according to the selected command.

For obstacle monitoring, an ultrasonic sensor continuously measures the distance between the robot and objects in its surrounding area. When an object comes within the specified distance limit, the controller can stop the motors to avoid a possible collision. A servo motor helps position the sensor in different directions for improved environmental monitoring. The speaker is used to provide audio indications or programmed responses during operation.

The integration of these hardware and software elements forms a robotic platform capable of wireless voice-command operation and basic obstacle detection.

V. SYSTEM ARCHITECTURE

A. General Architecture of the Proposed System

The Smart Voice-Based AI Robot is designed using a combination of electronic hardware and software components that function together to achieve voice-operated robotic movement. The architecture establishes communication between the user, the mobile application, and the robotic vehicle through a wireless interface.

Initially, the user provides a spoken instruction using a smartphone. The mobile application analyses the voice input and identifies the intended movement command. After recognition, the corresponding command is transferred to the ESP32-S3 development board through Bluetooth Low Energy (BLE).

The ESP32-S3 functions as the main controller and manages the operations of the connected modules. It examines the incoming command and sends appropriate signals to the motor driver. Based on these signals, the DC motors operate to move the robot in the selected direction. An ultrasonic sensor continuously checks the surrounding area to identify nearby obstacles. The servo mechanism changes the sensor's orientation to examine different directions, and the speaker produces audio feedback according to the programmed instructions.

The combined operation of these modules enables the robotic vehicle to respond to wireless voice commands and perform basic obstacle-monitoring functions.

B. Functional Description of Individual Modules

1. Voice Input Section: The voice input section is responsible for receiving instructions from the user. A smartphone is used to capture spoken information through its internal microphone. The captured audio is then supplied to the speech recognition facility of the mobile application for identification.

2. Mobile Application Section: This section acts as the communication interface between the user and the robot. It analyses the spoken instruction and

converts it into a command that can be understood by the controller. Common movement instructions, such as moving forward, reversing, and turning in either direction, are assigned corresponding command values before transmission.

3. Wireless Data Transfer Section: BLE technology is used to exchange information between the smartphone and the ESP32-S3 board. The recognized command is sent wirelessly to the robotic vehicle, allowing it to be operated remotely. This arrangement avoids the need for a direct physical connection during normal operation.

4. Controller and Processing Section: The ESP32-S3 is responsible for coordinating the complete robotic operation. It receives the command data from the smartphone and checks whether the instruction matches one of the programmed control functions. Once a valid command is identified, the board activates the appropriate output pins to operate the connected devices.

5. Motor Driving Section: The movement mechanism consists of a motor driver connected to the DC motors. The controller provides the required input signals to the driver, which controls the motors according to the selected movement instruction. By changing the motor operating directions, the robotic vehicle can move ahead, travel backwards, and perform left or right turns.

6. Distance Measurement and Obstacle Monitoring Section: The ultrasonic sensor is used to determine the presence of objects near the robotic vehicle. It emits high-frequency sound waves and detects the reflected waves returning from surrounding objects. The distance is calculated from the time difference between the transmitted and received signals. When an object enters the predefined detection range, the controller can interrupt motor operation and bring the robot to a stop.

7. Sensor Positioning Section: A servo motor is connected to the ultrasonic sensor to control its angular position. This arrangement allows the sensor to examine different directions instead of monitoring only one fixed position. The distance information collected at different angles can help the robot identify obstacles around its path.

8. Sound Feedback Section: The audio section consists of a speaker that communicates information to the user. Depending on the programmed functions, it can produce confirmation messages, warning sounds, or other predefined audio responses during the operation of the robot.

VI. ADVANTAGES

The proposed Smart Voice-Based AI Robot offers a convenient approach to controlling a robotic vehicle by combining voice recognition, wireless data transmission, and sensor-based operation. The important benefits of this system are explained below.

1. Voice-Activated Control: The robot responds to instructions provided through speech, allowing users to perform movement operations without relying on conventional manual controls.

2. Reduced Manual Interaction: The system minimizes the use of physical controllers because the required instructions can be provided directly through a smartphone.

3. Wireless Command Transmission: Bluetooth Low Energy provides a communication link between the mobile application and the ESP32-S3 board, enabling commands to reach the robot without cable connections.

4. Convenient Operating Method: The voice-based interface simplifies the process of interacting with the robotic vehicle and makes its operation more accessible.

5. Automatic Obstacle Monitoring: The ultrasonic module checks the distance between the robot and nearby objects. If an obstacle is detected within the programmed limit, the controller can deactivate the motors to help prevent collisions.

6. Controlled Navigation: The motor driver manages the rotation of the DC motors, enabling the robot to execute different directional movements based on the received instructions.

7. Audio Communication: The speaker provides audible information about selected operations and can deliver programmed notifications to the user.

8. Affordable Prototype Design: The use of readily available development boards and electronic modules makes the project suitable for educational institutions and small-scale prototype development.

9. Programmable Operation: The ESP32-S3 allows changes to be made to the control program, making it easier to adjust the robot's behaviour and introduce new operating instructions.

10. Possibility of Feature Enhancement: The existing arrangement can be upgraded by connecting additional sensing devices and communication modules, creating opportunities for more advanced robotic applications.

VII. KEY COMPONENTS

The proposed Smart Voice-Based AI Robot is developed using a combination of electronic hardware and software tools. Each component performs a specific function, and their coordinated operation enables voice-based navigation, wireless communication, and obstacle detection.

A. Hardware Components

1. ESP32-S3 Microcontroller: The ESP32-S3 serves as the primary controller of the robotic system. It receives movement instructions through BLE, processes the incoming data, and generates suitable signals to operate the connected devices.

2. Ultrasonic Sensor: This sensor is used to identify objects present near the robot. It calculates the distance to an object using ultrasonic wave transmission and reception, helping the system recognize possible obstacles.

3. Servo Motor: The servo motor changes the position of the ultrasonic sensor at different angles. This arrangement allows the robot to examine its surroundings from multiple directions.

4. Motor Driver: The motor driver acts as a power-control interface between the microcontroller and the DC motors. It receives signals from the ESP32-S3 and regulates motor operation according to the selected movement command.

5. DC Motors: These motors provide the mechanical force required for robot movement. By controlling their rotation, the system can perform forward, backward, left-turn, and right-turn operations.

6. Speaker: The speaker is used to deliver sound signals and pre-programmed voice notifications, providing audio information about the robot's activities.

7. Battery and Power Supply: The power unit supplies the required electrical energy to the entire circuit. It supports the operation of the controller, motor driver, sensors, and other electronic modules.

B. Software Components

1. Mobile Application: The mobile application serves as the communication interface between the user and the robot. It accepts voice input and converts the recognized instructions into commands that the robot can understand.

2. Speech Recognition: This software function analyzes the user's speech and determines the requested operation. The recognized instruction is then converted into a suitable command for controlling the robot.

3. Bluetooth Low Energy (BLE): BLE enables wireless data exchange between the smartphone and the ESP32-S3. It carries the selected control instructions to the robot without requiring a wired connection.

4. Arduino IDE: Arduino IDE is the software platform used to create and edit the robot's program. It also helps compile the source code and transfer the completed firmware to the ESP32-S3 board.

C. SUMMARY

The Smart Voice-Based AI Robot is intended to simplify robotic operation by allowing users to

control its movements through spoken instructions. The system uses a smartphone application to identify voice commands and translate them into specific movement instructions, including forward, backward, left, and right. The generated commands are communicated to the ESP32-S3 controller through Bluetooth Low Energy (BLE).

The ESP32-S3 acts as the main processing unit and manages the robot's operations based on the received instructions. It sends appropriate control signals to the motor driver, which regulates the connected DC motors to produce the required movement. For obstacle detection, an ultrasonic sensor measures the distance between the robot and nearby objects. If an object enters the predefined safety range, the controller can stop the motors. The servo motor changes the sensor's orientation to examine different areas, and the speaker produces programmed sound alerts or audio messages.

This project integrates microcontroller technology, wireless connectivity, and voice-controlled operation into a single robotic platform. It provides an alternative to conventional manual robot control and demonstrates the practical use of embedded systems in robotics.

The system operates through a sequence of activities. First, the user provides a spoken instruction through the smartphone. The mobile application processes the speech and identifies the corresponding command. The command is then transferred to the ESP32-S3 using BLE. Based on the received information, the controller activates the motor driver to operate the DC motors and perform the requested movement. At the same time, the ultrasonic sensor monitors nearby obstacles to support safer navigation.

Further development of the project may include more accurate voice recognition, improved navigation capabilities, advanced command interpretation, and additional safety functions.

VIII. CONCLUSION

The Smart Voice-Based AI Robot provides a voice-operated solution for controlling the movement of a robotic vehicle. In this system, a smartphone application recognizes the user's spoken instructions and converts them into predefined movement commands. These commands are transferred to the ESP32-S3 microcontroller through Bluetooth Low Energy (BLE). The controller then communicates with the motor driver to operate the DC motors and perform the required movements.

The system also includes an ultrasonic sensor for measuring the distance between the robot and nearby objects. When an obstacle is detected within the programmed distance, the robot can stop its movement to minimize the possibility of collision. The servo motor helps change the sensor's scanning angle, and the speaker is used to provide audio notifications during operation.

The project demonstrates the integration of voice-based control, embedded hardware, and wireless communication in a mobile robotic system. It offers a convenient alternative to traditional remote-control methods and provides a practical example of applying microcontroller technology to robotics. The developed model can be useful for academic demonstrations and basic robotic assistance applications.

The system has opportunities for further development. Future improvements may include better speech recognition in noisy surroundings, support for different languages, more accurate interpretation of user instructions, and improved autonomous navigation. Performance can also be assessed through experiments involving command recognition, communication delay, movement response, and obstacle detection. These enhancements could help improve the functionality and reliability of the robotic system.

VIII. REFERENCES

1. **Voice-Controlled Robotic Car**

Gupta, M., Tanushree, P., Konduru, V. K. R., & Eashanvi, A. (2025). “Design and Implementation of a Voice-Controlled Robotic Car for Real-Time Hands-Free Navigation and Control.” *2025 1st International Conference on Advancement in Futuristic Technologies (ICAFT)*. IEEE.

https://doi.org/10.1109/ICAFT66710.2025.11453137

2. **Voice-Based Robotic Manipulation**

Asha, C. S., & D'Souza, J. M. (2024). "Voice-Controlled Object Pick and Place for Collaborative Robots Employing the ROS2 Framework." *2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE).* IEEE. https://doi.org/10.1109/AMATHE61652.2024.10582225

3. **Speech Recognition in Human–Robot Interaction**

Goetzee, S., Mihhailov, K., Van De Laar, R., Baraka, K., & Hindriks, K. V. (2024). "Audio-Visual Speech Recognition for Human-Robot Interaction: A Feasibility Study." *2024 33rd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN),* pp. 930–935. IEEE. https://doi.org/10.1109/RO-MAN60168.2024.10731139

4. **Voice-Controlled Humanoid Robot**

Bharath Sai, U., Sivanagamani, K., Satish, B., & Ranga Rao, M. (2017). "Voice Controlled Humanoid Robot with Artificial Vision." *2017 International Conference on Trends in Electronics and Informatics (ICEI).* IEEE. https://doi.org/10.1109/ICOEI.2017.8300

5. **Ultrasonic Sensor-Based Robot Navigation**

Tripathi, A., Khan, M. A., Pandey, A., Yadav, P., et al. (2021). "Human Following Robot Using Ultrasonic Sensor." *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N).* IEEE.

https://doi.org/10.1109/ICAC3N53548.2021.9725768

6. **Speech Filtering for Human–Robot Interaction**

"A Near-Real-Time Processing Ego Speech Filtering Pipeline Designed for Speech Interruption During Human-Robot Interaction." (2024). *2024 33rd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN).* IEEE. https://doi.org/10.1109/RO-MAN60168.2024.10731262

7. **AI and Speech Processing in Human–Robot Interaction**

"Comparative Analysis of GPT-4 and LLaMA 3.2 Integration With Speech Processing Models for Enhancing Human–Robot Interaction and Motion Control in Real-World Applications." (2025). *IEEE Access,* vol. 13, pp. 127170–127182. IEEE. https://doi.org/10.1109/ACCESS.2025.3590592

8. **Verbal Commands for Robot Task Execution**

"Instructing Hierarchical Tasks to Robots by Verbal Commands." (2024). *2024 IEEE/SICE International Symposium on System Integration (SII).* IEEE. https://doi.org/10.1109/SII58957.2024.10417491

9. **Integrating AI Services into Conversational Robots**

"Integrating Multimodal AI Services into Pepper for Natural Human-Robot Interaction." (2025). *2025 Computing, Communications and IoT Applications (ComComAp).* IEEE. https://doi.org/10.1109/ComComAp68359.2025.11353191

10. **Voice Command Recognition and Sound Source Localization**

Korayem, M. H., Azargoshasb, S., Korayem, A. H., & Tabibian, S. (2021). "Design and Implementation of the Voice Command Recognition and the Sound Source Localization System for Human–Robot Interaction." *Robotica,* vol. 39, no. 10, pp. 1779–1790.

https://doi.org/10.1017/S0263574720001496