iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78066

Autonomous Robo Rider Operated by Hand Gesture with Obstacle Detection

Mrs. K. Raja Deepa[1], G. Naga Lakshmi[2], C. Guru Archana[3], P. Jaya Simha[4], B. Harshitha[5]
[1]*Assistant Professor,* [2, 3, 4, 5]*UG Students , Department of Electronics & Communication Engineering, Sai Rajeswari Institute of Technology, Proddatur, Andhra Pradesh-516360*

Abstract: This paper presents the design and implementation of a low cost gesture controlled robot using an ESP32 CAM based mobile platform with live video streaming, ultrasonic obstacle detection and wireless control via the ESP NOW protocol. The proposed System comprises an ESP8266/ESP12E based wearable transmitter unit equipped with an MPU 6050 inertial measurement unit (IMU) and an ESP32 CAM based receiver unit integrated with a differential drive robot chassis, LN98N motor driver, HC SR04 ultrasonic sensor and buzzer.
Gesture information acquired from the IMU is translated into motion commands and transmitted using ESP NOW to the ESP32 CAM, which simultaneously drives the motor, performs obstacle detection and streams real time video over a Wi Fi access point. The ESP NOW protocol is selected for its low latency, connectionless communication characteristics, which are advantageous for responsive robot tele operation compares to conventional TCP /IP or to end command latency below 40ms within a 10m line of sight indoor environment, while maintaining QVGA video streaming at 20 -24 frames per second. The integrated ultrasonic module detects frontal obstacles within 20cm and triggers an audible alert, enhancing operational safety.
Keywords: ESP32CAM, ESP8266 and ESP NOW gesture control, MPU6050, Ultrasonic sensor, obstacle detection, telepresence robot, mobile robotics, and wireless control.

I. INTRODUCTION

Mobile robots controlled in real time by natural human gestures are rapidly emerging as an attractive paradigm for telepresence, remote inspection, and educational robotics. Recent advance in low cost microcontrollers, integrated wireless connectivity and compact inertial sensors have enabled students and hobbyists to build sophisticated robotic platforms with capabilities that were previously limited to industrial or research laboratories. However, designing a system that concurrently supports low latency control, real time video feedback and reliable obstacle detection remains a challenging task, especially under the constraints of low power consumption, low cost, and limited computational resources.

The work presents in this paper addresses these challenges by realizing a compact gesture controlled telepresence robot based on the ESP32 CAM development board and an ESP8266/ESP12E transmitter. The ESP32 CAM integrates a dual core micro controller, Wi Fi connectivity and an OV2640 camera, thereby enabling image capture and HTTP video streaming without external processing units. In this work, the frontal environment, and when an obstacle is detected below a predefined threshold distance, an audible buzzer alert is activated, warning both the operator and surrounding users, this layered sensing approach enhances the safety and reliability of tele operated robots, especially in cluttered indoor spaces.

The main contributions of this paper are summarized as follows:

1) Design and implementation of a bi directional telepresence robot using an ESP32 CAM for simultaneous video streaming and motion control.
2) Development of a gesture based wearable transmitter using ESP8266/ ESP12E and MPU6050 for intuitive robot navigation.
3) Integration of HC SR04 ultrasonic sensor obstacle detection and buzzer alert into the motion control loop without degrading communication performance.

II. LITERATURE SURVEY

A. Gesture controlled Robots

Gesture controlled mobile robots have been widely explored using various sensing moralities including camera-based vision systems, wearable inertial sensors and capacitive or flex sensors. Camera-based sensors approaches typically utilize computer vision algorithms running on embedded processors or single board computers such as Raspberry pi, and interpret hand gestures and arm movements in image frames.

In contrast, wearable inertial sensors such as MPU6050, which combines a 3-axis gyroscope and 3-axis accelerometer, offer an efficient and low power alternative for gesture recognition. By directly measuring linear acceleration and angular velocity of the user's hand, mapping these signals to robot motion commands becomes straight forward and requires only simple thresholding or classification logic. [9]

B. ESP32 CAM Based Robots

The ESP32 CAM board has emerged as a popular platform for low cost wireless video streaming and basic computer vision tasks due to its integration of an ESP32 SOC and an OV2640 camera module. Project involving ESP32 CAM controlled robots often rely on HTTP web servers hosted on the device where users can view live streams and control motion via web based interfaces. While this architecture simplifies user's interaction, control commands are typically transmitted using higher level network protocols over TCP/IP, which are less suitable for time critical gesture control. Recent research on smart autonomous robots using ESP32 CAM has also focused on integrating ultrasonic sensors and object detection models for obstacle avoidance and environment perception.

These systems demonstrate the feasibility of combing video streaming with local sensing on a single ESP32 CAM, but they often emphasize autonomy rather that interactive teleoperation.

C. ESP NOW Based Wireless Control

ESP NOW is a connectionless wireless communication protocol introduced by espressif that enables peer to peer data exchange among ESP32 and ESP8266 devices using vendor specific action frames at the MAC layer. Unlike Wi Fi TCP/IP or UDP communication, ESP NOW does not require association with an access point, thereby reducing setup time and communication overhead. Measure latency for short payloads are typically in the order of a few milliseconds, which is attractive for real time control applications such as sensor network and robots.

Several works have demonstrated ESP NOW based communication between pairs or networks of ESP32/ESP8266 modules for environmental monitoring, home automation and simple remote control applications.

Table 1: comparison table

Comparison of Existing Approaches

Method	Technology Used	Advantages	Limitations
Gesture Controlled Robots	Vision systems and inertial sensors (MPU6050)	Natural human–robot interaction	High processing requirement and lighting dependency
ESP32-CAM Based Robots	Wi-Fi communication using HTTP/TCP-IP	Low cost with video streaming capability	Higher latency for real-time control
ESP-NOW Based Control	Peer-to-peer wireless protocol for ESP32	Very low latency and fast communication	Limited range and data payload

III. SYSTEM DESIGN

A. Overall System Description

The proposed system consist of two main subsystems: a wearable gesture transmitter and a mobile telepresence robot. The wearable transmitter is implemented on an ESP8266/ ESP 12E module and equipped with an MPU6050 IMU module .the telepresence robot is built around an ESP32 CAM module interfaced with an L298N dual H bridge motor bridge, two DC motors, an HC SR04 ultrasonic distance sensor and a piezoelectric buzzer.

The ESP8266 continuously reads acceleration data from the MPU6050 over the I^2C bus and computes normalized acceleration components along the X and Y axis. On the robot side, the ESP32 CAM receives gesture data, interprets it as direction command and drives the motor control pins accordingly. Simultaneously, the ESP32 CAM configures itself in AP+STA mode, creates a dedicated Wi Fi access point and runs an HTTP server to provide MJPEG video streaming from the attached camera.

B. Gesture Transmitter Subsystem

The gesture transmitter is realized using. An ESP8266 / ESP 12E development boarddpowered by a 5v USB source and regulated.to 3.3v.

Which are the recommended I^2C pins on Most ESP8266 / development boards. The ESP8266 operates in station mode without Association to any access point; instead, it Uses ESP NOW for direct communication.

The transmitter periodically samples the MPU6050 at a fixed rate (e.g., 20HZ). The acceleration readings along the X and Y axis, ax and ay, are normalized by the gravitational constant g to obtain dimensional values:

C. Telepresence Robot Subsystem

The telepresence robot subsystem uses an ESP32 CAM board boards as a central controller. Four GPIO pins (GPIO14, GPIO15, GPIO13 and GPIO12) are interfaced with the IN1-IN4 inputs of the L298N motor driver, implementing a differential drive configuration for forward, backward, left, right motions. The ESP32 CAM configures its camera using standard OV2640 settings, and MJPEG frames are served via an HTTP Server running on port80.

IV. MATHEMATICAL MODELING AND CONTROL ALGORITHMS

A. Gesture Mapping Model

The mapping from human hand tilt to robot motion is implemented using a threshold based state machine. Let a~x and a~y denote the normalized acceleration components along the forward-backward and left-right axis of the hand, respectively, as defined in (1). A symmetric dead zone threshold is applied to both components to filter out small, unintentional hand tremors.

This approach can be generalized to map continuous tilt values to proportional speed commands.in the current implementation, binary motion control is used, thus, these are quantized to discrete states corresponding to full forward, left rotation, right rotation and stop.

B. Ultrasonic Distance Measurement Model

The HCSR04 ultrasonic sensor emits a burst of ultrasonic waves at frequency f~40 KHZ and measures the time T for the echo to return from the nearest obstacle. Assuming the speed of sound in air is vs~343m/s at room temperature, the distance is computed.in

To meet a superior readings due to noise or multipath reflections, a simple temporal filtering scheme is employed. let dk denote the raw distance measurement at time step k. a moving average filter over an samples.in practice, a small window size n=3 or 5 is often sufficient smooth out noise while preserving responsiveness. The filtered distance d-k is then compared against the threshold dth to trigger the buzzer. [6]

C. ESP NOW Communication model

ESP NOW uses vendor specific action frames at the IEEE 802.11 MAC layer to transmit small players (up to 250 bytes) between peers without requiring establishment of a full Wi Fi connection. Where T queue is the time spent in the local transmit Q, TX is the actual frame transmission time over the air and tack corresponds to the optional acknowledgement delay. For the small payload sizes and low network congestion, Queue is negligible and TX is determined for primarily by the data rate and channel conditions. In experimental tests, it is common to observe ESP NOW one way latencies in the range of 2-10ms for short control messages. [1] [4]

V. BLOCK DIAGRAM

The entire system is implemented using the Arduino IDE with expressive ESP32 and ESP8266 board support packages. The ESP32 CAM firmware is structured in to separate modules: camera intelligence and HTTP streaming, ESP NOW receiver layer, motor control driver, ultrasonic measurement task and buzzer control logic. The main loop function primarily handles periodic ultrasonic measurements and leaves ESP NOW reception to an interrupt driven call back, thereby ensuring that gesture command are processed promptly even under continuous video streaming load.

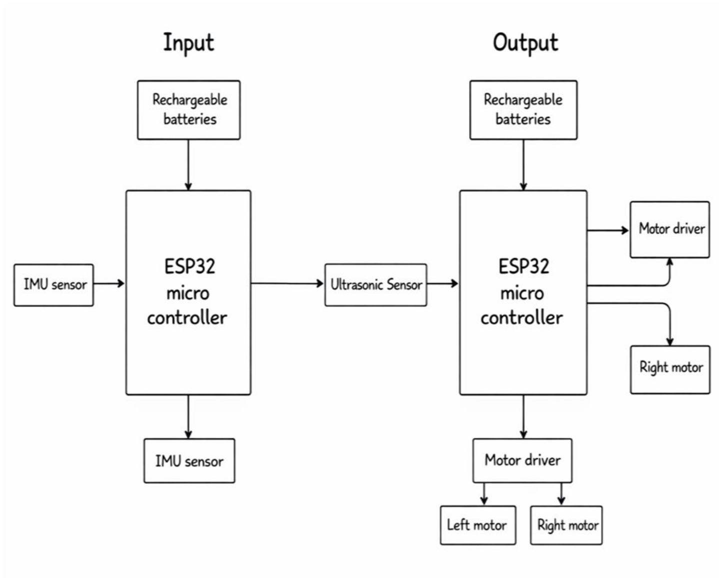


Fig 1: Block diagram of project

A. Components

1) ESP32 Microcontroller

Fig 2: ESP32 Microcontroller

ESP32 Microcontroller is a family is a energy efficient micro controllers that integrate both Wi-Fi and Bluetooth applications. It supports multiple GPIO pins, analog inputs, PWM, and serial communication. Due to its high performance and low cost, it is widely used in robotics and smart automation projects.

2) MPU 6050

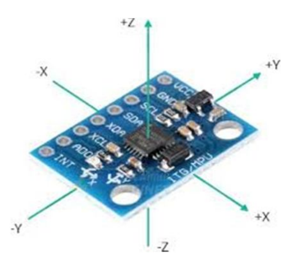


Fig 3: MPU6050 accelerometer

The MPU6050 devices combines a 3-axis gyroscope and 3- accelerometer which is. Processes complex 6-axis motion fusion Algorithms. It is used to measure tilt, acceleration, and angular rotation of the robot. In this project, the MPU6050 helps in detecting orientation and movement, improving stability and motion control of the robotic system. [5]

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

3) ESP32 CAM

Fig 4: ESP32 CAM

ESP32 CAM is a low-cost ESP32 based development board with on board camera. It is used for capturing and streaming real-time video over a wireless network. It enables remote surveillance and monitoring of the robot through a web interface. [2]

4) Motor Driver

Fig 5: motor driver

The driver plays a vital role in regulating the operations of electric motors, enabling precise control over their speed and direction. It acts as an interface between the microcontroller and motors, providing sufficient current and voltage for proper motor operation. The motor driver enables forward, backward, left, and right movement of the robot. [3]

5) BO Motor

Fig 6: BO motor

A Battery operated motor mainly consists of a DC motor, gear box, a plastic casing, and a shaft for wheel attachment. These motors are widely used in small robotic vehicles for efficient and smooth motion control. Measures the cloudiness of water to detect the impurity levels. Helps identify contamination by analysing suspended particles**.**

Implementation: The integration of multiple functional subsystems-gesture sensing wireless communication, motor control, video streaming and obstacle detection-on resource constraint microcontroller presents several implementation challenges. One key observation from the experiments is that careful scheduling and avoidance of blocking delays in the firmware are essential to maintain responsiveness. For example ultrasonic distance measurements must be executed with short timeouts to prevent blocking the main loop and delaying ESP NOW packet processing. From a human- robot interaction perspective the simple threshold based mapping used in this work is intuitive and easy to learn. Users can rapidly adapt to the tilt based control scheme, with forward and backward motions corresponding to role gestures. [7] [8] [10]

VI. RESULT

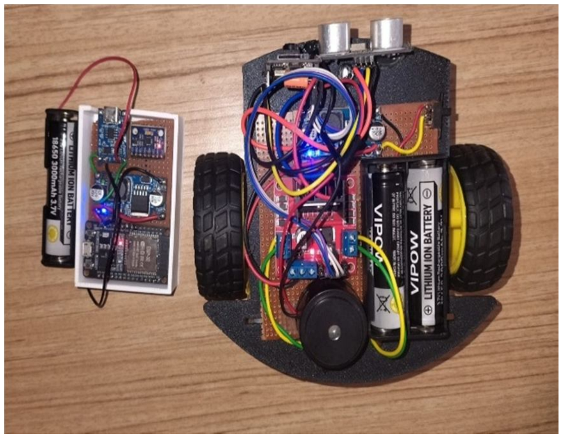


Fig 7: overall output

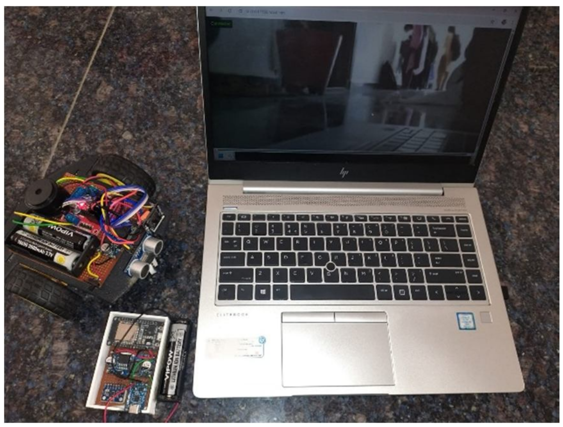


Fig 8: live streaming output

VII. CONCLUSION

This paper has presented the design, implementation and evaluation of a low cost gesture control telepresence robot biased on an ESP 32 CAM receiver and an ESP8266 /12E transmitter. The system integrates IMU based gestures sensing, ESP NOW low latency wireless communication, real-time video streaming, ultrasonic obstacle detection and audible alerts within a compact robotic platform. Experimental results demonstrate that the proposed architecture can achieve responsive control with low command latency and high packet delivery ration while maintaining acceptable video streaming performance in indoor environment.

The modular nature of the design makes it suitable for educational laboratories, project based learning and hobbyist robotics, while the combination of telepresence and obstacle detection opens avenues for applications in remote inspection, simple surveillance and human- robot interaction scenarios future work will investigate adaptive gesture recognition using embedded machine learning models, multi robot control using a single transmitter, integration of additional sensors such as infrared or time of flight modules, and partial of loading of video analytics to edge or cloud platforms.

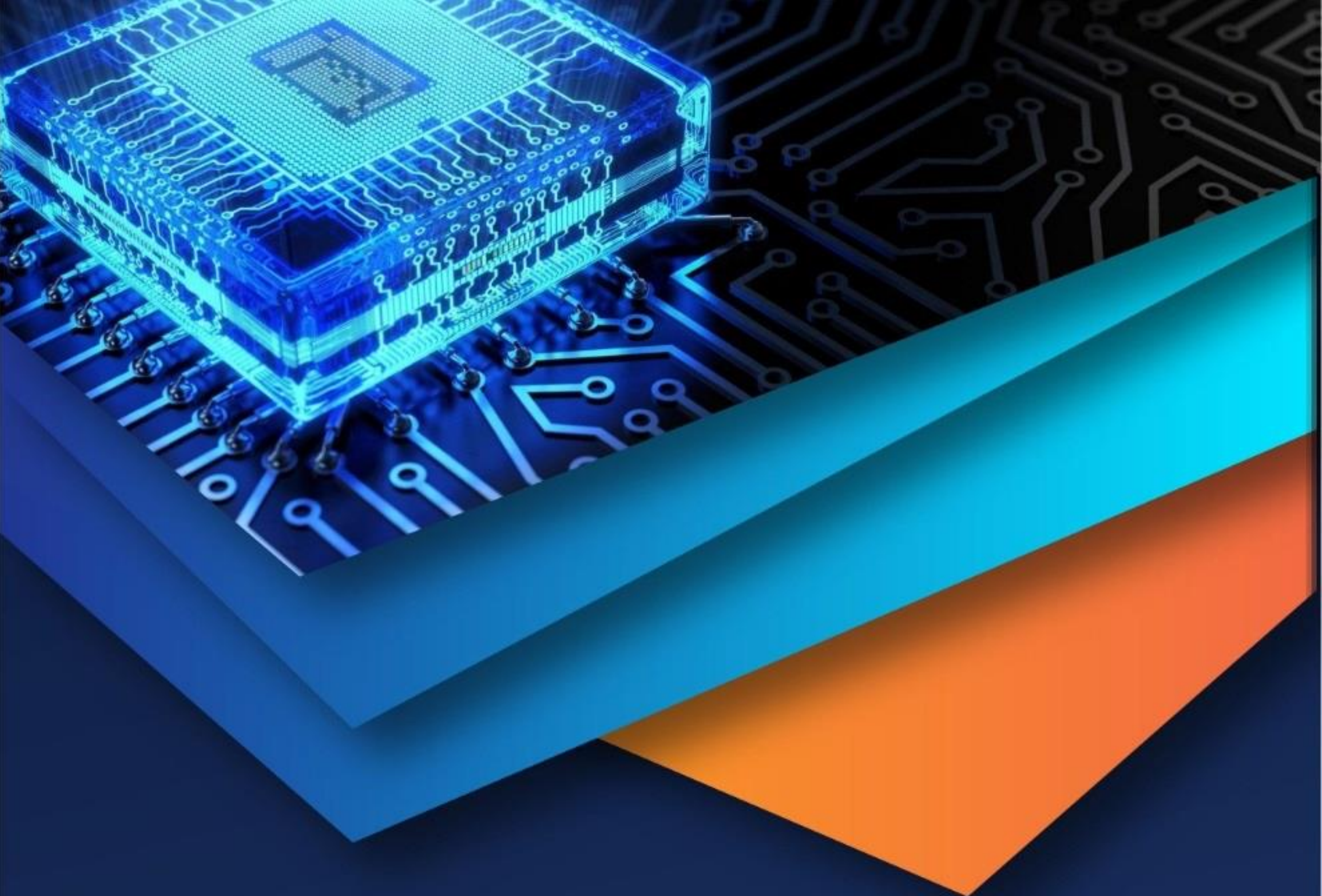
VIII. ACKNOWLEDGEMENT

The authors would like to thank department of electronics and communication engineering, Sai Rajeswari institute of technology for providing laboratory facilities and support during the development of this project the authors also express their gratitude to their project guide continuous technical guidance and motivation throughout the work.

REFERENCES

[1] Espressif Systems, “ESP-NOW User Guide for ESP8266 and ESP32,” Espressif Documentation, 2023.
[2] Espressif Systems, “ESP32-CAM Technical Reference Manual,” Espressif Systems, 2022.
[3] Random Nerd Tutorials, “ESP32-CAM AI Thinker Pinout Guide: GPIOs Usage Explained,” 2025.

[4] Random Nerd Tutorials, "Getting Started with ESP-NOW Using ESP8266 NodeMCU," 2024.
[5] InvenSense, "MPU-6050 Six-Axis Motion Tracking Device Datasheet," TDK-InvenSense, 2021.
[6] SparkFun Electronics, "HC-SR04 Ultrasonic Distance Sensor Datasheet and Application Guide," 2022.
[7] R. Kumar and S. Patel, "Wi-Fi Controlled Robot Using ESP32-CAM with Live Video Streaming," *International Research Journal of Engineering and Technology*, vol. 11, no. 6, pp. 2145-2149, 2024.
[8] A. Sharma and P. Gupta, "Smart Autonomous Robot with Obstacle Detection," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 12, no. 3, pp. 1123-1128, 2025.
[9] J. Park and H. Kim, "Gesture-Based Robot Control Using Inertial Sensors," *IEEE Sensors Journal*, vol. 20, no. 14, pp. 7890-7897, 2020.



10.22214/IJRASET

INDEX COPERNICUS 45.98

ISRA JIF IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY