Artificial Intelligence Powered Hexa-Wheeled Robot

[1]Birva A. Dave, [2]Brijraj R. Kacha

[1,2] U.G. Student, Department of Computer Engineering, Atmiya University, Rajkot, Gujarat, India.

***Abstract*— "Victor Uniform 2" (VU2) is an Artificial Intelligence (AI) powered mobile robot with six wheels which is developed for different purposes. It has developed functions such as speech recognition, gesture sensitivity and live broadcasting capability. This research analyses the technologies used in the design and the assembly of VU2. The study also deals with the identification of the problems which arose in the course of the development and the ways in which they were solved.**

Keywords—Robot, Artificial Intelligence, ESP32, Gemini, Gesture Control, Hexa-Wheel

I. Introduction

The increasing need for efficiency, accuracy, security or advanced functions in processes such as surveillance, quarantine, hospitality, exploration and others. This report concerns with the design of AI-powered six wheeled robot 'VU2' (Victor Uniform 2). While developing the VU2, the concept of a Mars rover was used as its concept design. To work on the idea of mars rover design, a study carried out by Firat Barlas on "Design of a Mars Rover Suspension Mechanism" 2004 [1] is taken as a reference. This system allows the robot to be controlled the by means of movement of the hand in a special glove. The robot has been equipped with double lambda suspension mechanism cutting off about two inches of frames from lower to upper frame making it easily move over small barriers. It can work for a time span of 1.2 hours on simple standby and a wireless connection with an operational range of 100 meters and distance (these figures can be changed to suit operational needs).

The robot can interface with humans using voice. For this purpose, it possesses a speech recognition and response system which is integrated with the Google Cloud Platform API (Application Programming Interface) and the Gemini API. The robot consists of an Arduino Microcontroller for Motion, an Arduino Nano for Robot control, an ESP32 for speech recognition and response, and an ESP32 AI Thinker CAM for live streaming of the video over a local area network. As for the robotics applications, they cut across sectors such as research and development, hospitality, hospitals, and industries. Moreover, scrap plastic is used to make the robot's body; hence it helps to deal with waste and meets sustainable development objectives. Tests have been done in detail to test the accuracy and effectiveness of the robot.

II. Literature Review

A. *Hand Gesture Controlled Robot*

Using hand gestures as a means of control improves the interaction with the robot as it is more comfortable and easier than the traditional methods of controlling the desktop robots using the keyboard or a joystick. It also improves the way a user's hand is tracked and recognized in order to aid and/or improve user interactivity as well as to provide a finer touch for the task that commands it. This step increases the overall user satisfaction by ensuring that the task is completed efficiently, within the limitations of what is achievable.

Numerous articles have been dedicated to the problem of resolving robots' addressing without any support through hand movements. In 2019, Bakri et al. presented an experiment on a hand-controlled robotic arm with gesture-recognition module (NRF24L01) [2]. In 2020, Vadlamudi et al. designed a robot controlled by hand movements with Arduino and MPU6050 [3]. In 2023, Amalia Hasibuan et al. described the development of a wheeled robot controlled by hand gestures based on MPU-6050, L298N motor driver, and Arduino [4]. One more strong reason for an additional investigation and development of the concept of control of robotic arms by hands is based on studies described above.

B. *Surveillance using ESP32 Cam*

The most distinctive feature of the ESP32-CAM is the presence of an integrated camera along with Microcontroller ESP32 so that video images can be captured and transmitted over a network in real-time. Remote camera monitoring and security surveillance is the main benefit. A recent study by Deshpande et al. (2024) explored the possibility of the implementation of the ESP32-CAM in the design of surveillance mobile robots [5].

The study showed that the module was capable of recording and streaming live video content, accessing Wi-Fi, and customizing features as required. These features make it possible to use the ESP32-CAM for the development of effective solutions for security and monitoring purposes.

C. *Rocker Bogie Mechanism*

Rocker-bogie designs are especially suitable for robots which operate in hard to traverse environments. Study carried out by Abhisek Verma and his team about "Design of Rocker-Bogie Mechanism" in May 2017 [6], provides insights on how this mechanism performance on obstacle clearance and improves stability due to its advanced suspension system. The study helped to develop a similar mechanism for the robot. Having this mechanism incorporated using connected rocker arms and bogies allows for all six wheels to always be in contact with the ground even when the robot comes across obstacles more than one wheel's diameter thereby providing support and movement so that the robot does not seize. This is further achieved due to the independent rocker arms since the robot is capable of adjusting to the terrain and therefore has less pitching motion and improved stability.

D. The Google Cloud API and Gemini API

Speech and text elements are embedded in modern machines, useful in enhancing the interaction between humans and machines.

According to Iancu (2019) [7], the Google Cloud Speech to Text API has made it possible for Iancu to connect spoken commands to robot comprehension in real time. Another similar, but more powerful, API is Google Cloud Gemini which is able to formulate a text.

With the use of these APIs robots can be able to listen, speak, and interact with people at a much higher level than before.

III. Methodology

A. Electronic Hardware Components

1) Microcontrollers: Arduino Uno R3, Arduino Nano, ESP32 Dev Module, ESP32 AI Thinker CAM Module.

2) Sensors, Actuators, and Other Modules: MPU6050 6-axis(combines 3-axis Gyroscope, 3-axis Accelerometer) motion tracking device, L298N Motor Driver, nRF24L01 Transceiver, IR (Infra-Red) Sensor, INMP441 Microphone, MAX98357A Amplifier, 3W 4Ω Speaker, DC (Direct Current) 12V Battery.

B. Base Technology

1) Google Cloud Speech-to-Text API: The API can recognize the speech that corresponds to the user's commands.

2) Gemini API: The API allows you to send text and the system will respond with appropriate text.

3) C++ Language: Programming language used for firmware development.

IV. Detailed Explanation Of The Robot

VU2 is a robot with six wheels and a double-lambda suspension, which also has capabilities of speech recognition and response as well as live video streaming. It consists of three main sections:

1) *The Body:* Responsible for movement and navigation.
2) *The Communicator:* Allows conversing with the user.
3) *The Eye:* Supplies sight as well as live video feed.

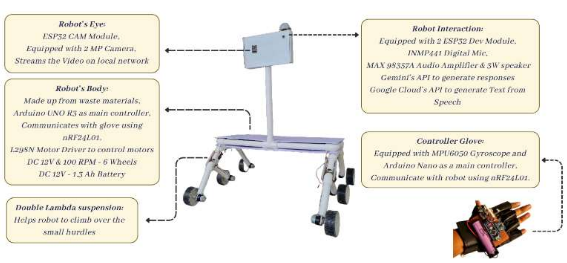

Fig. 1. VU2 Robot's features with it's Gesture control hand glove.

A. 'The Body'

The Body can be further divided into 2 parts. The Navigator, is the gesture control glove that controls the movements of the robot. And another is the Actuator, which follow the commands from the Navigator and providing movement to the robot.

1) The Navigator: The Navigator controls robot movements, including an Arduino Nano, MPU6050 sensor, nRF24L01 module, and 3.3V battery. The battery powers the glove, the MPU6050 senses gestures, and the nRF24L01 communicates with the Actuator. The Arduino Nano integrates these components.

2) The Actuator: The Actuator controls robot's neck, body, legs, and wheels. The Arduino Uno communicates with the Navigator using nRF24L01. The L298N motor driver controls 6 gear motors. It is powered by a 12V DC battery.

3) Materials and Design: Materials used include PVC (Poly Vinyl Chloride) pipes, wooden strips, and aluminium for construction. The double lambda suspension was inspired by Mars rovers.

4) Power Supply: Power Supply is a 12V DC lead-acid battery with 1.3 Ah capacity, providing one hour of runtime. It can be replaced with a higher capacity battery.

5) Working: User wears the glove and turns it on, followed by the robot. The gyroscope calibrates for 30 seconds. Hand movements control the robot's movement. MPU6050 sensor detects hand positions and transmits data to Arduino Nano via I2C (Inter Integrated Circuit) protocol. Arduino Nano makes movement decisions and transmits them to the Actuator using nRF24L01 transceiver modules. Actuator receives data and transmits it to Arduino Uno R3 controller via SPI (Serial Peripheral Interface) protocol communication. Arduino Uno R3 controls the 6 gear motors using the L298N motor driver module. The circuit diagrams for the Navigator and Actuator components are presented below.

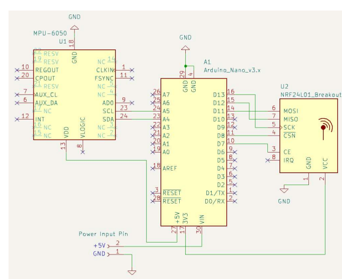

Fig. 2. Circuit Diagram for the Navigator (Hand Glove).

The circuit should operate with an input voltage (V-in) within the range of 4V to 9V. A 3.3V cell when fully charged would deliver around 4.2V DC which falls within the acceptable input voltage range, hence the transmitter will work well.

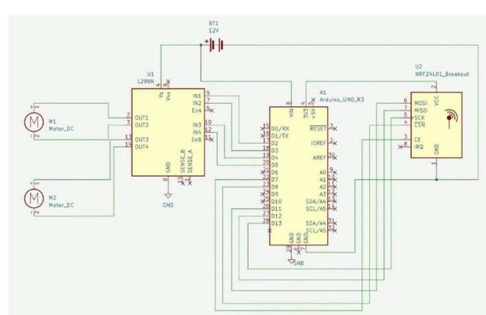

Fig. 3. Circuit Diagram for the Actuator (Robot Body).

The diagram above represents the reference circuit for the Receiver (Rx) code. In the reference circuit, only two motors are employed. The implementation connects six motors. Additionally, while the reference design utilizes a 9V battery to power the system, whereas a 12V DC battery is used in this application.

B. *'The Communicator'*

Contains two ESP32s, an IR sensor, INMP441 digital microphone, MAX98357A audio amplifier, and speakers. Users activate the microphone by placing a hand over the IR sensor. Activated microphone stays active for five seconds. The INMP441 microphone can also convert sound signals into digital data which is done through the use of the I2S (Integrated Inter-IC Sound) library. The digital data is then sent to Google's Cloud Speech to text API. Produced text is fed into Google's Gemini API. The text in response is then formatted to an audio response using the Audio library of MAX98357A and relayed back to the user through the speaker. Two ESP32 microcontrollers efficiently control and run the system. One ESP32 incorporates the IR sensor, INMP441 microphone and Speech to Text API. The other ESP32 incorporates the Gemini API, the audio amplifier and loudspeaker. Both the ESP32 communicate with each other through their Universal Asynchronous Receiver-Transmitter (UART) protocol.

The circuit diagram for this communicator is presented below. The design of the communicator component was guided by instructional materials from the Techiesms channel on YouTube [8].

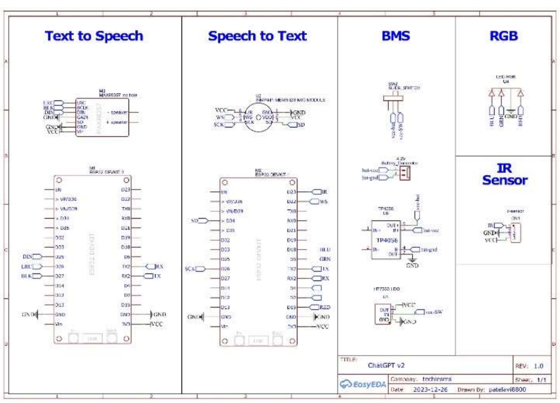

Fig. 4. Schematic Diagram for the Communicator (Credits: techisms).

C. *'The Eye'*

ESP32 AI Thinker Cam Module enables live video streaming over a local network. It has a 2-megapixel camera and flash for enhanced image quality. The video streaming is accessible only on a local private network. It requires a 5V DC input, obtained by using a 7805-voltage regulator Integrated Circuit (IC) and capacitors to convert the 12V DC power source. The firmware for the module is uploaded using an FTDI (Future Technology Devices International Limited) programmer chip. To stream video, the firmware code must include the Wi-Fi network name and password. Upon successful upload, the module will return a private IP (Internet Protocol) address in the Arduino IDE's (Integrated Development Environment) serial monitor. To watch the video stream, connect the viewing device to the same Wi-Fi network as the ESP32 AI Thinker CAM module and enter the private IP address in a web browser.

Fig. 5. Implementation of live video streaming in VU2 Robot.

V. EXPERIMENTS & TEST RESULTS ON THE BODY

A. *Structural Design*

An iterative design process was used in the course of improving the double lambda suspension and PVC pipes were used for cost-effectiveness and sustainability. The size of wheels selected was 10 centimeters of diameter due to the budget and design conditions. The size of the robot was reduced by half making it 2:1 in length to width ratio, in order to improve maneuverability of the robot. Testing the robot has included intensive usability assessments such as stair descent. It was necessary to carry out minor design works for improving the turning ability of the model. The last dimensions included; 3 feet wide, 2 feet 4 inches longer, and 3 feet 6 inches tall.

Fig. 6. Body design of the latest version of the Robot.

B. Power Supply

The initial plan to use Li-ion batteries for the power supply encountered challenges during testing. For these reasons, a lead-acid battery was used as it offered better performance. Changing to the lead-acid battery assured that there would be a steady power supply for the robot.

C. Response Generation

Initially, OpenAI's ChatGPT API was employed for response generation. However, due to rising costs, the model was transitioned to Gemini, a more affordable option. Using Gemini enabled reduction on costs due an initial allotment of free tokens, which made it a better option for the project.

VI. Results & Discussion

The robot was created at an affordable rate and tested substantively. It proved to have ruggedness, over an hour working time on battery, and a distance range of about 100 meters extendable to 1 kilometer. The skeleton of the robot is designed from recyclable substances. Its activities address six Sustainable Development Goals (SDGs) — 3,4,8,11,9 and 17. It can be used in a number of areas such as education, health, industry, and hospitality. Its features can be improved in the subsequent research.

VII. Future Enhancement

The future studies are directed towards advancement of robotic design, deepening the accuracy, enhancing power efficiency, and increasing the working radius. Integrating new materials and technologies, researchers would be able to design a more robust, more responsive, and more enduring robot. In particular, further scenarios could include the followings.

A. Design Optimization

Developing a more robust and aesthetically pleasing design to enhance durability and appeal.

B. Precision and Accuracy

Improving the robot's ability to provide precise and accurate responses to user commands and environmental stimuli.

C. Power Efficiency

Exploring techniques to increase the robot's power efficiency and extend its operational time.

D. Operational Range

Investigating methods to increase the robot's operational range and allow it to be deployed in larger areas.

Research on these areas would further improve the current abilities of the robot, as well as advance its uses in other fields.

VIII. Conclusion

The complex design and simple intuitive use of the six wheeled rover robot VU2 demonstrates the practical application of robotics in everyday life. Its design is ingeniously conceived by bearing in mind the design principles of the Mars Rover. VU2 offers user-friendly gesture control and robust construction for navigating challenging terrains. Its live video streaming capabilities enable remote monitoring, hence pointing out the eases automation and artificial intelligence can bring to such problems.

The success of VU2 lays the groundwork for future developments in robotics, demonstrating the possibility of efficient machines that are both highly versatile and user friendly, that could easily fit into human environments. The integration of gesture control and live video streaming offers valuable insights for enhancing robot interactivity, hence facilitating creative ways of problem solving and development.

Acknowledgement

Authors of this paper are thankful to Dr. Ashish Kothari, Director Research Innovation & Translation, Atmiya University, Rajkot, Gujarat, India and Ms. Tosal Bhalodia, Head Department of the Computer Engineering, Atmiya University, Rajkot, Gujarat, India for providing experimental resources and support. Some of the coding insights are taken from the YouTube Channel named “techiesms” operated by Mr. Sachin Soni.

References

[1] Firat Barlas, “Design of a Mars Rover Suspension Mechanism” M.S. dissertation, Department of Mechanical Engineering, Izmir Institute of Technology, Izmir, Turkey, June 2004.

[2] A. B. Bakri, R. Adnan and F. A. Ruslan, "Wireless Hand Gesture Controlled Robotic Arm Via NRF24L01 Transceiver," 2019 IEEE 9th Symposium on Computer Applications & Industrial Electronics (ISCAIE), Malaysia, 2019, pp. 16-22, doi: 10.1109/ISCAIE.2019.8743772.

[3] Vadlamudi, Syambabu & Kumar, D. & Kumar, G.. (2020). Hand Gesture Controlled Robot using Arduino and MPU6050. International Journal of Recent Technology and Engineering (IJRTE). 9. 777-779. 10.35940/ijrte.D9546.059120.

[4] Hasibuan, Azhari & Nasution, T & Azis, P. (2023). MPU-6050 Wheeled Robot Controlled Hand Gesture Using L298N Driver Based on Arduino. Journal of Physics: Conference Series. 2421. 012022. 10.1088/1742-6596/2421/1/012022.

[5] Hrishikesh Deshpande, Dnyaneshwari Zite, Sakshi bagul and Dr. Vineeta Philip, “RESEARCH PAPER ON SURVEILLANCE ROBOT USING ESP 32 CAM” in International Journal of Creative Rsearch Thoughts (IJCRT) – 2024. [Online] Available: https://www.ijcrt.org/papers/IJCRT24A4148.pdf

[6] Abhisek Verma, Chandrajeet Yadav, Bandana Singh , Arpit Gupta, Jaya Mishra, Abhishek Saxena, “Design of Rocker-Bogie Mechanism” in International Journal of Innovative Science and Research Technology – 2017, Volume 2, Issue 5, May – 2017. [Online] Available: https://ijisrt.com/wp-content/uploads/2017/05/Design-of-Rocker-Bogie-Mechanism-1.pdf

[7] Iancu, Bogdan. (2019). Evaluating Google Speech-to-Text API's Performance for Romanian e-Learning Resources. Informatica Economica. 23. 17-25. 10.12948/issn14531305/23.1.2019.02.

[8] Techiesms, S. Soni, Gujarat, India. *“ChatGPT on ESP32”* (2023). Accessed: January, 2024. [Online Playlist]. Available: https://youtube.com/playlist?list=PLruzZCuhcsGOzuC0RAPJoh5bIcJIHHHAk&feature=shared