INTERNATIONAL JOURNAL OF ADVANCED INNOVATIVE TECHNOLOGY IN ENGINEERING

Published by Instigative Research Private Limited Home page: www.ijaite.co.in

IoT-Based Voice and Gesture-Controlled Smart Home

[1]Prof. Akshata Kumbhare, [2]Kunal Kawale, [3]Limesh Naikar

[1,2,3]Department of computer Engineering, SCET, Nagpur, Maharashtra, India

[1]akshatakumbhare26@gmail.com, [2]kunalkawale7890@gmail.com, [3]limeshnaikwar@gmail.com

Article History

Received on: 20 July 2026

Revised on: 15 Aug 2026

Accepted on: 20 Sept 2026

Keywords: Gest Control, Voice Recognition , human -Computer interaction, Metapipe Hands , Speech Recognition

e-ISSN: 2455-6491

DOI: 10.65809/IJAITE/26/v03sp02/031

Production and hosted by
www.insightiveinc.org

©2026|All right reserved.

ABSTRACT

This paper presents the design and implementation of a home automation system that is controlled using both voice and gesture commands. It integrates Internet of Things (IoT) technology with human-computer interaction techniques to develop an intelligent, user-friendly, and accessible solution. The system provides real-time control of home appliances using technologies such as Raspberry Pi, Media Pipe, and MQTT protocol, ensuring low latency and secure communication. Users can manage devices like lights, fans, and security systems through simple voice instructions and hand gestures, enabling a completely hands-free experience. Experimental results indicate that the combined use of voice and gesture control offers greater flexibility, accessibility, and reliability compared to traditional single-mode systems. The study highlights that multimodal interaction plays a key role in building more inclusive and efficient smart home environments. IoT integration ensures seamless communication between smart devices, central control units, and cloud. allowing remote access and monitoring. Additionally, smart energy management techniques are incorporated to track device.

1. INTRODUCTION

Home automation has changed a lot in the last ten years. It has turned regular homes into smart spaces that can do things with little help from people. Artificial Intelligence (AI), the Internet of Things (IoT), and Natural Language Processing (NLP) are some of the advanced technologies that have made it possible to control home appliances with voice commands and hand gestures. This study aims to create a smart home automation system that uses both voice and gesture recognition to make it easier for people to use, more efficient, and convenient. Voice-controlled systems let users talk to devices instead of using physical switches or remote controls. Gesture recognition lets people control appliances with simple movements, which is especially helpful for people who have trouble moving around.

Combining these two interaction methods provides a flexible and user-friendly interface, accommodating diverse user preferences and environmental conditions. The proposed system aims to bridge the gap between traditional automation systems and next-generation smart environments by integrating cost-effective hardware components such as microcontrollers, sensors, and cameras with intelligent software algorithms.

The system processes user inputs in real-time and translates them into actionable commands for controlling appliances such as

lights, fans, and electronic devices. Furthermore, the growing demand for energy-efficient solutions has increased the importance of smart systems that optimize energy consumption. By enabling precise control and automation, the system contributes to reducing energy wastage and improving sustainability. This study underscores the significance of multimodal interaction in smart homes and illustrates how the integration of voice and gesture controls can markedly improve usability, reliability, and accessibility, facilitating future progress in human-computer interaction.

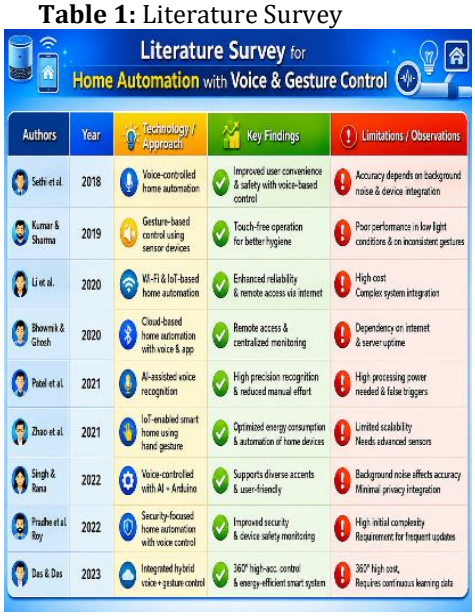
2. LITERATURE SURVEY

Home Automation as a topic has received significant attention in recent years, with several more researchers proposing different methods for home automation including IoT-based systems, voice recognition systems, and gestconsystems. Traditional home automation systems is totally depended mainly on wired communications d and simple remote controls, making them less flexible and scalable. The introduction of wireless communication, however, such Some researches have concentrated on designing voice-controlled home automation systems is the through the use of voice recognition software or speech recognition engines like Amazon Alexa, Google Home, etc. Such systems use cloud computing to recognize the voice commands of the users and carry out specific actions. Although accurate in their results, they require an active internet connection to perform their tasks effectively. Gesture control is another interface technology that has received much interest owing to its intuitiveness. Computer vision and sensor-based approaches have been employed in recognizing gestures. The former has great flexibility, although its overall performance can be affected by light and background. Sensor-based gesture recognition systems, which utilize accelerometers and infrared sensors, on The other hand, are more reliable, but they might limit user mobility.

Voice and gesture control technologies can be combined to take advantage of their respective strengths while mitigating their drawbacks. Such hybrid systems offer redundancies such that hold one method fail, there is always another means available. Research indicates that multimodal interfaces enhance user satisfaction and accessibility, especially for older people and those with disabilities.

However, despite all these developments, there is still much to improve regarding the cost of the system, its accuracy, efficiency, and reliability. The majority of existing solutions are very costly and complicated to be used in less economically advanced countries. This paper, therefore, aims to develop on previous studies by suggesting a new cost-efficient solution.

Table 1: Literature Survey



3. PROBLEM DEFINITION

Home automation systems traditionally used manual switches or single-mode interactions that are not only cumbersome but difficult for some users. The use of applications that are entirely reliant on mobile phones for their operation is not always convenient and might be difficult to use by people such as the aged and disabled due to the fact that these applications cannot function offline Another limitation associated with current home automation systems is the exorbitant prices that limit its accessibility and wide acceptance. The devices used need a constant supply of internet connection and may be expensive for installation in other regions apart from the rich industrialized nations.

The issue of security and privacy remains another major challenge. Voice-activated devices mostly utilize cloud computing, which means that there might be cases of data leakage or intrusion by unauthorized personnel. In addition, the problem of errors when interpreting the commands issued by users may also cause some problems and create additional complications. At present, there is a lack of a simple interface that could take into account the individual characteristics of users. The existing technologies do not cater for the needs of people with disabilities who difficulty using the standard control buttons or even smartphones.

In this regard, there is a need to develop a robust home automation system that incorporates various modes of interaction, minimizes the dependence on internet connection, and provides accurate recognition with enhanced levels of safety.

4. OBJECTIVES

The main objective of this study is the development of an intelligent home automation system that uses voice and gestures as modes of

controlling the system in order to enhance user experience. In the process, this will lead to less user effort as far as controlling devices is concerned.

One of the important objectives of this system is the achievement of high recognition accuracy when it comes to voice commands and gesture recognition. This will involve the use of sophisticated techniques that will increase the speed and accuracy of the system.Another very important objective of this project is to design a cost-effective system that can be adopted by a majority of people in the residential setting. Affordable hardware parts and free software will go a long way in ensuring this.

Apart from sustainability, the study will also focus on making the use of energy more efficient through an effective way of controlling and monitoring the use of appliances. Through the use of automated control devices, the wastage of energy could be minimized and a sustainable lifestyle could be achieved.

Furthermore, the study will seek to ensure security and privacy through reduced use of cloud-based operations and provision of secure communication protocols. Protecting data and restricting access to data forms part of the main objectives of the proposed study.

Lastly, there is need for scalability of the project. This means that apart from fulfilling the core objectives of the study, it should also be flexible enough to incorporate other components like mobile operation.

5. PROPOSED WORK

The system under consideration is a multichannel home automation system that employs both gesture and voice recognition methods for the effective operation. This system was developed with the goal of improving upon existing solutions that are restricted to using either gestures or voices as an input channel, providing enhanced accuracy and reliability.

This home automation system contains three basic elements: input module, processor module, and control/output module. An input module comprises a microphone for receiving voice commands from the operator and a camera or sensor for recognizing gestures made by hands. Voice input data undergo speech recognition in order to be converted into text-based command strings. The resulting strings are compared against a library of commands stored in a database of the system.

For the purpose of gesture recognition, the camera module and image processing algorithms are employed. Real-time processing of the video stream captured by the camera allows identifying hand gestures. Methods like background subtraction, contour detection, and pattern recognition are employed for identification of gestures, which are mapped onto appropriate control commands.

A processing unit (microcontroller or a single-board computer) is responsible for controlling the entire process. The unit processes the information obtained from both input modules by applying decision-making algorithms. An algorithm is employed to resolve any conflicts when multiple inputs occur at once so that there would be no ambiguity regarding which command will be executed.

An output module of the system includes relay switches connected to electrical appliances in a house like lights, fans, etc. According to the command obtained as a result of the processing, signals are sent to relays that activate or deactivate the appliance. Feedback mechanisms may be included into a system for notifying users about executing their commands.

Error handling is considered to increase system dependability. Thus, in case of failed voice command because of noise, one could use gestures as an alternative means of control. Another measure is confirmation prompts in order to avoid accidental actions.

Cost-efficiency and scalability of the system are achieved due to usage of affordable hardware parts and software. Scalability implies further integration of more devices and capabilities. Some security measures are added to increase protection: basic authentication and local data processing.

Thus, it is possible to conclude that this system is a flexible solution for home automation in the modern era.



Figure 2: Proposed work

6. METHODOLOGY

The suggested system consists of hardware and software elements that integrate to facilitate control of home appliances via voice and gestures. The hardware elements include a microcontroller, sensors, camera modules, and communication mediums like Wi-Fi or Bluetooth.

Voice recognition is facilitated by means of software modules that process sound waves produced by the speaker, which are subsequently

mapped onto predefined functions. The processing process involves noise elimination, feature extraction, and pattern matching. It employs offline processing techniques in order to minimize internet dependency.

Gesture recognition is carried out with the help of a camera or motion sensors. Computer vision techniques like image processing are used for recognizing certain gestures. It is determined on the basis of the movement of the hands by the system.

The two input types are incorporated within a control unit where the commands are analysed, and the signals are sent out to the respective appliances to perform the tasks assigned. Relays are used to turn off or on the appliances based on the received signals.

There is a feedback component in the system in order to notify the user of the outcome of their command. Feedbacks can be delivered via sound, graphics, and mobile alerts. Several experiments were performed under varying conditions to assess the efficiency and performance of the systems. The tests included checking for accuracy, delay in processing, and reliability.

In conclusion, the methodology aimed at developing an efficient and reliable system of controlling various appliances using voice and gestures.

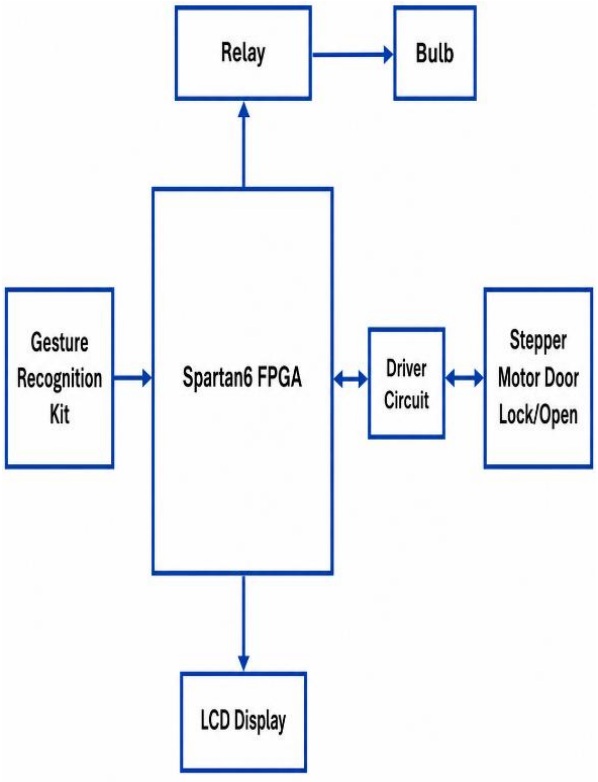


Figure 3: Sentiment Analysis Visualization

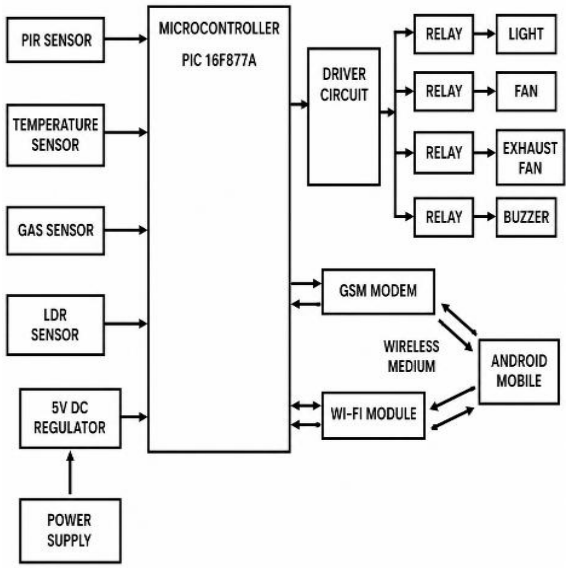


Figure 4: Block Diagram for Home Automation

CLASSICIFICATION

The home automation system may be classified depending on their interaction process, means of communication, and intelligence type. Depending on the interaction process, systems are classified as manual, voice-based, gesture-based, or multimodal. The proposed system is a multimodal system.

Depending on communication mode, systems are either wired or wireless. The proposed system is a wireless system offering more convenience and scalability. Depending on the system architecture, systems are either centralized or distributed. The proposed system adopts a centralized system architecture.

Depending on the intelligence level, systems vary from automated to intelligent. Since the system involves machine learning, it is an intelligent system.

CONCLUSION

This study illustrates the design and development of an intelligent IoT-based automated home system with voice and gesture recognition technology. The proposed system overcomes the shortcomings of conventional systems by incorporating multiple modes of interaction and user-friendly features. The presented system showcases the efficiency of incorporating multiple modes like voice commands and gestures into an IoT setup. This is a cost-effective, scalable, and secure option for smart further studies could involve the use of machine learning and other AI technologies for prediction automation and more efficient security and device connective

REFERENCES

[1] M. Kumar, et al. (2017) – "Gesture Based Smart Home Automation Using Arduino," IEEE Transactions on Automation Science and Engineering

[2] S. Patel, et al. (2019) – "Voice Controlled Home Automation Using Bluetooth Module," IJSER, Vol. 10, Issue 3.
[3] Sharma, et al. (2020) – "IoT-based Smart Home Automation System," IJERT, Vol. 9, Issue 6.
[4] K. Singh, et al. (2022) – "Integration of Gesture and Voice Control in Smart Homes," Springer Journal of Intelligent Systems.
[5] R. Gupta & T. Verma (2021) – "IoT Enabled Home Automation with Voice and Gesture Control," International Research Journal of Engineering and Technology (IRJET).
[6] P. Jain & S. Yadav (2020) – "Design and Implementation of Smart Home using NodeMCU," IEEE Xplore Conference on IoT Applications.
[7] N. Ahmed, et al. (2018) – "Low-cost Home Automation Using Arduino and Bluetooth," IJCSIT, Vol. 9, Issue 5.
[8] R. Piyare & M. Tazil (2017) – "Bluetooth Based Home Automation System Using Cell Phone," IEEE Consumer Electronics Journal.
[9] K. Sharma & R. Kumar (2021) – "Gesture Controlled Smart Devices using MPU6050," International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE).
[10] T. K. Roy & S. Banerjee (2023) – "Smart Home Automation Using Voice Assistant and IoT Cloud Integration," IEEE Access, Vol. 11.