



INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

Impact factor: 4.295

(Volume 5, Issue 1)

Available online at: www.ijariit.com

Smart virtual assistant bot using voice control and object detection

Abhishek B. R.

abhishek.regina@gmail.com

Sri Krishna College of Technology, Coimbatore,
Tamil Nadu

Gokul Krishnan V.

linkwithgokul@gmail.com

Sri Krishna College of Technology, Coimbatore,
Tamil Nadu

Balamurugan S.

bmurugan035@gmail.com

Sri Krishna College of Technology, Coimbatore,
Tamil Nadu

Anitha R.

r.anitha@skct.edu.in

Sri Krishna College of Technology, Coimbatore,
Tamil Nadu

ABSTRACT

Manual efforts being put up by human beings are greatly reduced by smart robots. The robot is voice-controlled and uses object detection using a virtual assistant. The human voice commands are given to the robotic assistant remotely by using a smart mobile phone or a raspberry pi microcontroller or a computer via the virtual assistant. The robot can perform different movements such as turn, start/stop operations, displaying the temperature and humidity of the environment. This robot also locates an object when moving from one place to another. The voice commands are processed in real-time, using an online cloud server. The speech signal commands converted to text form are communicated to the robot over a Wi-Fi network. This robot contains one Raspberry Pi and NodeMCU microcontroller as the central unit. The camera is connected to Raspberry Pi for the detection of objects that is then streamed to the users mobile or computer. All the motors and sensors are connected to the NodeMCU microcontroller. This idea is used for surveillance purpose, healthcare assistant and human-machine interaction purposes.

Keywords— Object detection, Raspberry Pi, NodeMCU, Virtual assistant

1. INTRODUCTION

A smart robot is an artificial intelligence (AI) system that can learn from its environment and its experience and build on its capabilities based on that knowledge. Smart robots can collaborate with humans, working along-side them, learning from their behaviour and have the capacity for not only manual labour but cognitive tasks. Smart robots help human beings in reducing manual efforts in day-to-day tasks and the risk to precious human lives in hazardous situations. They are typically used in a place where human contact is not possible.

1.1 Voice control

Voice or speaker recognition is the ability of a machine or program to receive and interpret dictation or to understand and

carry out spoken commands. Voice recognition has gained prominence and uses with the rise of AI and intelligent assistants, such as Amazon's Alexa, Apple's Siri and Microsoft's Cortana. The first thing that is considered taking about voice control is Speech Recognition i.e., making the system to understand human speech. Speech Recognition is a technology where the computer understands the words given through the natural language of speech. Speech is an ideal method for controlling robots and in communication. Speech recognition works using algorithms through acoustic and language modelling. Acoustic modelling represents the relationship between linguistic units of speech and audio signals; language modelling matches sounds with word sequences to help distinguish between words that sound similar. Voice recognition systems enable consumers to interact with technology simply by speaking to it, enabling hands-free requests, reminders and other simple tasks. Voice recognition software on computers requires that analog audio is converted into digital signals, known as analog-to-digital conversion. For a computer to decipher a signal, it must have a digital database, or vocabulary, of words or syllables, as well as a speedy means for comparing this data to signals. The speech patterns are stored on the hard drive and loaded into memory when the program is run. A comparator checks these stored patterns against the output of the A/D converter which is an action called pattern recognition. The size of a voice recognition program's effective vocabulary is directly related to the random-access memory capacity of the computer in which it is installed. A voice recognition program runs many times faster if the entire vocabulary can be loaded into RAM, as compared with searching the hard drive for some of the matches. Processing speed is critical, as well, because it affects how fast the computer can search the RAM for matches.

1.2 Use of building robots

Robots are indispensable in many manufacturing industries. The reason is that the cost per hour to operate a robot is a fraction of the cost of the human labour needed to perform the

same function. More than this, once programmed, robots repeatedly perform functions with a high accuracy that surpasses that of the most experienced human operator. Human operators are, however, far more versatile. Humans can switch job tasks easily. Robots are built and programmed to be job specific. You wouldn't be able to program a welding robot to start counting parts in a bin. Robots are in the infancy stage of their evolution. As robots evolve, they will become more versatile, emulating the human capacity and ability to switch job tasks easily. While the personal computer has made an indelible mark on society, the personal robot hasn't made an appearance. Obviously, there's more to a personal robot than a personal computer. Robots require a combination of elements to be effective sophistication of intelligence, movement, mobility, navigation, and purpose. Without risking human life or limb, robots can replace humans in some hazardous duty service. Robots can work in all types of polluted environments, chemical as well as nuclear. They can work in environments so hazardous that an unprotected human would quickly die.

1.3 Object detection

An image classification or image recognition model simply detect the probability of an object in an image. In contrast to this, object localization refers to identifying the location of an object in the image. An object localization algorithm will output the coordinates of the location of an object with respect to the image. In computer vision, the most popular way to localize an object in an image is to represent its location with the help of bounding boxes. Object recognition is widely used in the machine vision industry for inspection, registration and manipulation tasks. In many industries, robot arms need a mechanism to recognize objects to act on them, in an autonomous way, but the algorithms for object recognition have many limitations because of the changes in illumination, occlusion, scales and positions. Because of these limitations, there are many research groups working on projects to provide efficient solutions for object recognition in many fields, and robotics is the one where this project will be focused on. Since object recognition will be used on a robot, choosing an adequate software and hardware platform is a key factor. As many projects are being developed at universities, open-source software and cheap hardware have become a trend because it lowers the costs.

1.4 Virtual assistant

Virtual Assistants have a much-sophisticated interactive platform. They understand not just the language but also the meaning of what the user is saying. They can learn from instances and provide an unpredictability to their behaviour. That way they can have a long human interaction. They can be set to perform slightly complicated tasks as well. There has been a substantial amount of research in NLP to build advanced capabilities in virtual assistants; case in point Virtual Assistants can now understand slangs used in everyday natural conversations and analyse the sentiments by the use of languages, to enhance an even better set of communication skills. The NLP makes virtual assistants more conversational than chatbots. Virtual assistants can handle conversations, have robust NLP capacities and carry out a limited number of conversations via hard-coding, wildcard matching of words and time-consuming keyword training. Virtual Assistants have a wider scope and can perform a range of tasks, for example comparing products or finding the best product based on the given features. It is also deployed to tasks like decision making and e-commerce. It can perform activities like sharing jokes, playing music, stock market updates and even controlling the

electronic gadgets in the room. Unlike chatbots, virtual assistants mature gradually with use. Once people feel comfortable with voice search and in particular talking to smart speakers, the options become almost endless. By nature, our queries become longer and we want to have conversations when we're using smart speakers. As this technology improves and the quality of the answers to our questions become better and more reliable, this will become even truer. But in the meantime, the bottom line is people use their virtual assistants for all kinds of stuff, whether on a mobile device or via the smart speaker.

2. LITERATURE SURVEY

Rajnikant P, Sandhani and students explain about the smart robot which is made to monitor and check the situation of your home life from a remote place through the internet by google assistant and artificial intelligence. The paper published by Kumar and Chauhan explains about a robot that works with the input given a voice command. The system consists of three parts i.e., voice recognition system, a central controller system and the robot itself. The voice recognition system acts as an interface between the man and the robot. The robot receives the voice as an input signal and processes it and forward to the central controller and finally transmit the signal to the robot to execute the command. Emad S. Othman, states the implementation of a Voice Command System as an Intelligent Personal Assistant (IPA) that can perform numerous tasks or services for an individual. Raj Kumar Mistri focusses on Raspberry Pi board utilizing OpenCV and TensorFlow packages to perform actions. Ananthapadmanabhan J, Ms Annu Mariam Abraham, Mr Libin M George, Ms Vineetha Anna Saji, and Anil A explains the use of Arduino to control a smart bot. The survey is presented in a structured way, centred around three generations of algorithms: autonomous, cooperating, and variable structure. It emphasizes the underpinning of each algorithm and covers various issues in algorithm design, application, and performance. Generally, these kinds of systems are known as Speech Controlled Automation Systems (SCAS). The system is the prototype of the same. The idea is to develop some sort of menu driven control for the robot, where the menu is going to be voice driven and is controlled by a microcontroller remotely using the internet and the cloud servers. This system provided only the remote control through the internet and there was no usage of object detection through virtual assistant employed.

2.1 Existing system

This uses the Raspberry Pi 3 module for object detection controlled using a smartphone. There are no sensors for temperature and humidity measurement of the environment that includes real-time monitoring of the weather. A wireless module that sends data over the internet using Arduino micro-controller. The processing time is slow because of the relatively weak processing unit of Arduino UNO to control the rover movement.

2.2 Proposed system

The robot can perform basic 5 commands, left, right, go, back, forward, detecting temperature and humidity along with detecting objects in real time to interact with the user. NodeMCU controller-based voice recognition method is displayed in a smartphone. Using IFTTT (If This Then That) we create our own applets and connect them with the Adafruit cloud server. A trigger is created and then connected to an action to perform our specific task. In this case, the trigger service will be 'Google Assistant' and the Action will be

implemented through this service. The input from the user is through Google Assistant. The raspberry pi module is used for object detection in real time and is controlled by the user's phone or computer. This is much faster than the already used Arduino microcontroller that has relatively slow processing speed.

3. SYSTEM METHODOLOGY

The following are the modules in the system:

1. Voice Recognition Integration
2. Robot Movement and Sensor Usage
3. Object Detection Phase
4. Interactive Virtual Assistant

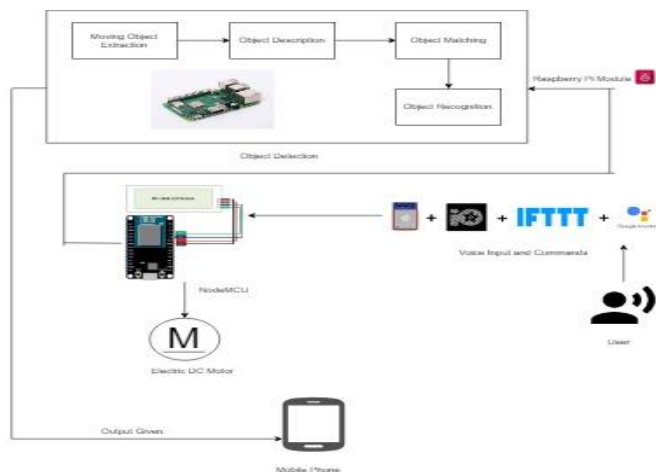


Fig. 1: Block diagram of proposed system

3.1 Voice recognition system

In this module, we focus on integrating the google assistant as a trigger to the actions on the IFTTT (If This Then That) platform through the Adafruit cloud server. The Adafruit cloud server consists of the data from the device which is displayed to the user. Each of the commands through the Google Assistant invokes an action specified in the IFTTT platform such as "Move forward 400" specifies the action to move the rover 400units in the forward direction.

3.2 Robot movement and sensor usage

In this module, the movement of the robot based on the commands from the server is executed. The temperature and humidity of the area are also displayed in the LCD by the DHT11 sensor. Each of the movements is executed by the driver shield that is connected to the NodeMCU unit. This unit is supplied 5V of power supply from the NodeMCU microcontroller that is sufficient for the whole circuit. The movements are also displayed on the 16x2 LCD display unit each time when a command is issued and when the rover performs the action.

3.3 Object detection phase

In this phase, the program for object detection is employed. The program is written in python language using open CV library and tensor flow. For detecting the object, we can give the object to the bot in two ways first direct via programming, second by the camera. In this case, the datasets are preloaded on the file and the extraction is made from the preloaded images. The goal

is to construct a model that can recognize the protest of indicated shading that make utilization of open source equipment and that chips away at the premise of visual information caught from an ordinary camera which has a reasonable lucidity. The proposed calculation is executed on the Raspberry Pi board utilizing OpenCV and Tensor flow.

3.4 Interactive virtual assistant

Alexa is also integrated into the raspberry pi module which uses speech to answer queries and respond to tasks. A speaker is connected to the module to enable audio streaming. Dynamic interactivity enabled.

4. CONCLUSION

This robot has multiple features which will give more security to any place. The temperature, humidity and real-time object detection is made possible and viewed on a mobile phone and computer. There are varieties of sensors for analyzing the current situation, a lot more can also be added according to the need. Everything can be controlled and monitored using a mobile or a computer through the internet from any part of the world. The user can taste the latest artificial intelligence technology that makes life a lot more secure and comfortable than before. The Voice Command System has enormous scope in the future. Like Siri, Google Now and Cortana become popular in the mobile industry. This makes the transition smooth to a complete voice command system. It's all about that frictionless experience that makes our lives much easier.

5. REFERENCES

- [1] Wai-Mo Mo Khaing "WIRELESS CONTROLLED SPY ROBOT USING PIC MICROCONTROLLER", International Journal of Science, Engineering and Technology Research, Volume 3, Issue 9, September 2014.
- [2] Sandesh Hathwar B, Manasa M B, Fathima Afshan and Chaithra R, "Raspberry Pi and IoT based Family Robot", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 6, Issue 7, July 2017.
- [3] A study on real-time control of a mobile robot with based on voice command, Byoung-Kyun Shim; Yoo-Ki Cho; Jong-Baem Won; Sung-Hyun Han Control, Automation and Systems (ICCAS), 2011 11th International Conference on Publication Year: 2011
- [4] Speech recognition and its application in the voice-based robot control system, Luo Zhizeng; ZhaoJingbing Intelligent Mechatronics and Automation, 2004. Proceedings. 2004 International Conference,
- [5] A study on precise control of an autonomous driving robot by voice recognition, Sung-Won Jung; Ki-Won Sung; Moon-Youl Park; Eon-Uck Kang; Won-Jun Hwang; Jong-Dae Won; Woo-Song Lee; Sung-Hyun Han Robotics (ISR), 2013 44th International, Symposium on DOI:10.1109/ISR.2013.6695640 Publication Year: 2013
- [6] A mechanism for human-robot interaction through informal voice commands, D'Mello, S.; Lee McCauley, Markham, James Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop, DOI:10.1109/ROMAN.2005.1513777, Publication Year: 2005.