EasyChair Preprint

№ 3681

Voice Control Human Assistance Robot

Linda John, Nilesh Vishwakarma and Rajat Sharma

June 27, 2020

Proceedings of the 2nd VNC; VNC-2020;
2020 2nd National Conference on "Technical Advancements for Social Upliftment", 04th April, 2020
Vidyavardhini's College of Engineering & Technology (VCET), Vasai (INDIA)

Voice Control Human Assistance Robot

Linda John
Project Guide (I.T.)
St. John College of Engineering and
Management
Mumbai, India
lindamaryj70@gmail.com

Nilesh Vishwakarma
Final Year (I.T.)
St. John College of Engineering and
Management
Mumbai, India
nilesh061294@gmail.com

Rajat Sharma
Final Year (I.T.)
St. John College of Engineering and
Management
Mumbai, India
rajatsharma1127@gmail.com

Abstract— **In today's world technology has made our lives easier and more enjoyable for the people. However, we all benefit from this emerging technology, certain groups of people need more help and support than others: elderly or disable people. For them, technology can make a way of having an almost normal life. So, we focused our attention on the concept of a personal assistant robot. This automaton is principally designed for this cluster of individuals as its main purpose is to supply help to Associate in Nursing aged or disabled persons.**

Personal robotic assistants help to reduce the manual efforts being put by humans in their day-to-day tasks. The purpose is to implement a voice-controlled system as an Intelligent Personal Assistant (IPA) that can perform numerous tasks or services for an individual. This golem is specially designed for this cluster of individuals as its main purpose is to supply help to associate senior or disabled person.

The human voice command is given to the robotic assistant remotely, by using a voice module (it is like ears of the robot). The tasks are based on some features embedded in the Assistant. The automaton will perform different movements, turns, start/stop operations. The robot can also read and recognize characters like alphabets as well as numeric. This enhances the capability of the robot to detect objects and relocate them from one place to another. The most important characteristic of the Assistant is to give the trending information that the user has asked for. The input query by the user can be news regarding weather, politics, etc.

Keywords — Voice Recognition, Assistant, Robot, Command.

I. INTRODUCTION

Over the years, humans have evolved in investing new technologies for reducing human efforts and saving human life [4]. Physically challenged and elderly people face difficulties while handling objects and hence they need an assistant for the same. Thus, if a robotic assistant is developed which can be operated using speech commands. It should be able to perform sure tasks we have a tendency to set for it. The desired task should be achieved among some given limitations. There are several robots are used by industries for easy and fast working. It may be human-controlled or automated.

There is a unit growing demands on artificial intelligence and automation within the service sectors for several years because of the age quake, silver society, and manpower shortage. Now that robotics is thus advanced with detector, mechanism and process intelligence at the side of the CPU development, it is expected to make much wider applications to the service sector and also for the personal assistant robot.

Our main objective is to design a less costly and effective personal assistance robot. This can do some small jobs for humans like move any object from one place to another. Basically, it is designed to carry some small weights and it can provide information about things that are asked by the user with the help of the internet. This robot will fetch the information form provided URLs (USING APIs) like Wikipedia. This Robot will be automated using machine learning algorithms to do things in particular areas like home. Also, the robot will find its way by using obstacle detection and character reorganization using ultrasonic sensors. All of the above features will be done by using the Raspberry Pi. New technologies based on speech, object and face recognition have become a complementary system for disabled people. Usually, they convert the human environment into speech or tactile information. Blind people or people with low vision may perceive persons from the environment, familiars, friends or S colleagues at work by face detection and recognition systems. Real-time object detection face recognition, text recognition, and currency bills identification are some of a large number of developed applications.[4],[6],[7]

So here a Robot is developed such that it has all the above-mentioned features of Face and object recognition which gives speech output such that it can interact with the Humans. The system is implemented on Raspberry Pi hardware. Raspberry Pi cams have a free open code and are able to run under OpenCV libraries and C++ bindings for Python. The Raspberry Pi 5 Megapixel camera is used to capture an image.

OpenCV termed as Opensource Computer vision is the latest and popular library started by Intel in 1999. OpenCV is a particular software that is used especially for image processing

in real-time. There are different versions in OpenCV and from the version, OpenCV 2.4 comes with Face Recognizer class for face recognition, such that suitable versions for Face recognition. The robot can be controlled by using a voice. Speech recognition is the process of converting speech to digital data, voice recognition is aimed toward identifying the person who is speaking. Voice recognition works by analyzing the features of speech that differ between individuals. So, by this, the movement of the robot can be controlled by using the voice to do things as like a human which can be used in many fields especially for blind people [4]

II. RELATED WORK

In 2015, Anurag Mishra, Pooja Makula, Akshay Kumar, Krit Karan and V. K. Mittal [1] This paper describes easy and simple hardware for implementation of Face, Object and speech detection and recognition. using an online cloud server. The speech signal commands converted to text form are communicated to the robot over a Bluetooth network.

In 2012, Xianghua Fan, Fuyou Zhang, Haixia Wang, and Xiao Lu [2] Face detection technology has widely attracted attention due to its enormous application value and market potential, such as face recognition and video surveillance system.

Real-time face detection not only is one part of the automatic face recognition system but also is developing an independent research subject. So, there are many approaches to solve face detection. Here the modified AdaBoost algorithm based on OpenCV is presented, and experiments of real-time face detection are also given through two methods of timer and dual-thread. The result shows that the method of face detection with dual-thread is simpler, smoother and more precise.

In 2018, Dyah Ayu Anggreini Tuasikal, Hanif Fakhrurroja, Carmadi Machbub [3] This paper describes Voice activation speaker recognition to regulate the Bioloid GP automaton by MFCC and DTW strategies is enforced well in automaton robots. The first step in the speech recognition process is feature extraction. In this paper use, Mel Frequency Cepstrum Coefficient (MFCC) on the characteristic extraction process and Dynamic Time Warping (DTW) used as a feature matching technique.

In 2017, U Bharath Sai, K Sivanagamani, B Satish, UG Students[4]. Today's most advanced industrial robots will soon become "dinosaurs." 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 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 labor 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 [4].
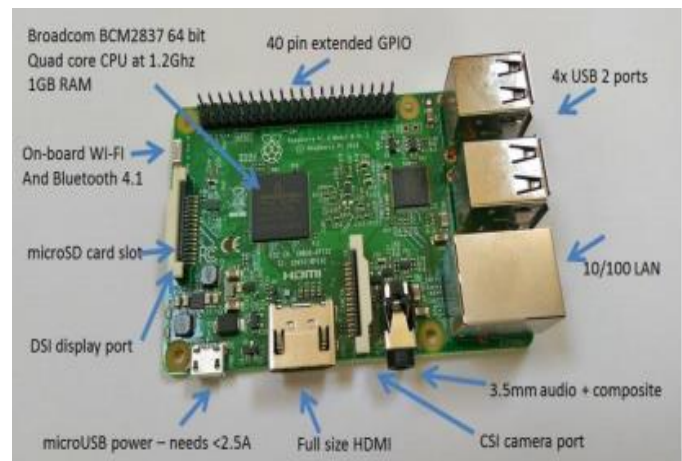
III. HARDWARE

A. RASPBERRY PI:

Fig-1: Raspberry pi Board

The Raspberry Pi 3 Model B+ is the latest product in the Raspberry Pi 3 range, boasting a 64-bit quad-core processor running at 1.4GHz, dual-band 2.4GHz and 5GHz wireless LAN, Bluetooth 4.2/BLE, faster Ethernet, and PoE capability via a separate PoE HAT The dual-band wireless LAN comes with modular compliance certification, allowing the board to be designed into end products with significantly reduced wireless LAN compliance testing, improving both cost and time to market.

The Raspberry Pi 3 Model B+ maintains the same mechanical footprint as both the Raspberry Pi 2 Model B and the Raspberry Pi 3 Model B.A powerful feature of the Raspberry Pi is the row of GPIO (general-purpose input/output) pins along the top edge of the board. A 40-pin GPIO header is found on all current Raspberry Pi boards (unpopulated on Pi Zero and Pi Zero W). Prior to the Pi 1 Model B+ (2014), boards comprised a shorter 26-pin header [8]. Any of the GPIO pins can be designated (in software) as an input or output pin and used for a wide range of purposes.
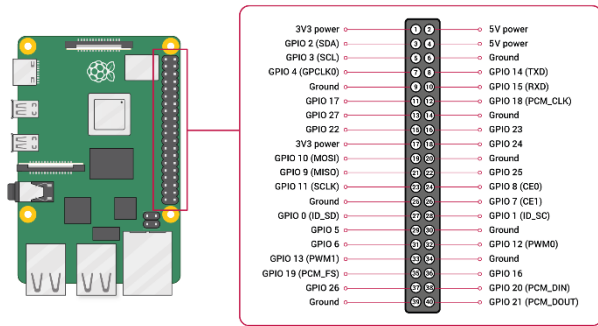
Fig-1.1: Raspberry pi Pins

Any of the GPIO pins can be designated (in software) as an input or output pin and used for a wide range of purposes.
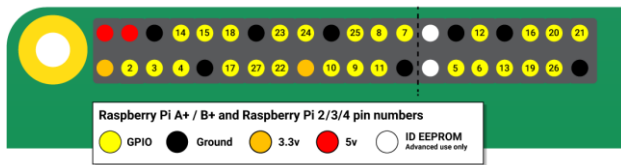

Fig-1.2: Raspberry pi Pins

Note: the numbering of the GPIO pins is not in numerical order; GPIO pins 0 and 1 are present on the board (physical pins 27 and 28) but are reserved for advanced use (see below).

Voltages
Two 5V pins and two 3V3 pins are present on the board, as well as a number of ground pins (0V), which are unconfigurable. The remaining pins are all general-purpose 3V3 pins, meaning outputs are set to 3V3 and inputs are 3V3-tolerant.

Outputs
A GPIO pin designated as an output pin can be set to high (3V3) or low (0V).

Inputs
A GPIO pin designated as an input pin can be read as high (3V3) or low (0V). This is made easier with the use of internal pull-up or pull-down resistors. Pins GPIO2 and GPIO3 have fixed pull-up resistors, but for other pins, this can be configured in software.

More
As well as simple input and output devices, the GPIO pins can be used with a variety of alternative functions, some are available on all pins, others on specific pins.

PWM (pulse-width modulation)
Software PWM available on all pins
Hardware PWM available on GPIO12, GPIO13, GPIO18, GPIO19

SPI
SPI0: MOSI (GPIO10); MISO (GPIO9); SCLK (GPIO11); CE0 (GPIO8), CE1 (GPIO7)
SPI1: MOSI (GPIO20); MISO (GPIO19); SCLK (GPIO21); CE0 (GPIO18); CE1 (GPIO17); CE2 (GPIO16)
I2C
Data: (GPIO2); Clock (GPIO3)
EEPROM Data: (GPIO0); EEPROM Clock (GPIO1)
Serial
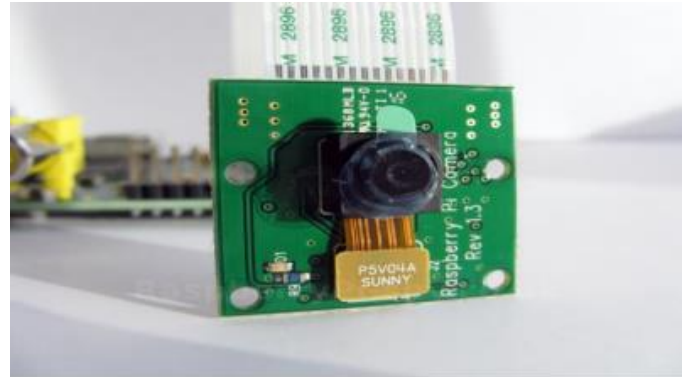TX (GPIO14); RX (GPIO15).

A. PI-CAMERA:

Fig-2: Pi Camera

The above-shown figure is the Camera module of Raspberry pi camera which is an official product from Raspberry pi. We are using camera module 1 which has technical specifications of 5 Megapixel camera of resolution 2592*1944 and weighs around 3 grams which have an Omni vision OV5647 Sensor. We interface this sensor with Raspberry pi for image processing. In an autonomous robot, it needs to perceive its environment through sensors in order to make logical decisions on how to act in the world. One important sensor in a robot is using a camera. There are different types of high-end cameras that would be great for robots like a stereo camera, but for the purpose of introducing the basics, we are just using a simple cheap webcam or the built-in cameras on our laptops.

B. MOTOR DRIVER

Fig-3: Dc Moto Driver L298n

This L298 Based Motor Driver Module is a high power motor driver perfect for driving DC Motors and Stepper Motors. It uses the popular L298 motor driver IC and has the onboard 5V regulator which it can supply to an external circuit. It can control up to 4 DC motors, or 2 DC motors with directional and speed control

This motor driver is perfect for robotics and mechatronics projects and perfect for controlling motors from microcontrollers, switches, relays, etc. Perfect for driving DC and Stepper motors for micro mouse, line-following robots, robot arms, etc.

An H-Bridge is a circuit that can drive a current in either polarity and be controlled by Pulse Width Modulation (PWM).

Pulse Width Modulation is a means of controlling the duration of an electronic pulse. In motors try to imagine the brush as a water wheel and electrons as the flowing droplets of water. The voltage would be the water flowing over the wheel at a constant rate, the more water flowing the higher the voltage. Motors are rated at certain voltages and can be damaged if the voltage is applied too heavily or if it is dropped quickly to slow the motor down. Thus PWM.

Take the water wheel analogy and think of the water hitting it in pulses but at a constant flow. The longer the pulses the faster the wheel will turn, the shorter the pulses, the slower the water wheel will turn. Motors will last much longer and be more reliable if controlled through PWM.

Motor A truth table			
ENA	IN1	IN2	Description
0	N/A	N/A	Motor A is off
1	0	0	Motor A is stopped (brakes)
1	0	1	Motor A is on and turning backward
1	1	0	Motor A is on and turning forwards
1	1	1	Motor A is stopped (brakes)

Motor B truth table			
ENB	IN3	IN4	Description
0	N/A	N/A	Motor B is off
1	0	0	Motor B is stopped (brakes)
1	0	1	Motor B is on and turning backward
1	1	0	Motor B is on and turning forwards
1	1	1	Motor B is stopped (brakes)

IV. PROPOSED SYSTEM

The proposed system summarizes the given input which can be in text or voice form. The voice input undergoes a voice to text conversion and is then applied to the summarization process where the importance of words is calculated based on word embedding and a new sentence is formed. In such a way a summary is generated and is given as an output. The most advanced version of the system can be used to generate minutes of the meeting. In this, the voice of the speaker is taken as input and this input is further processed to produce minutes of the meeting.
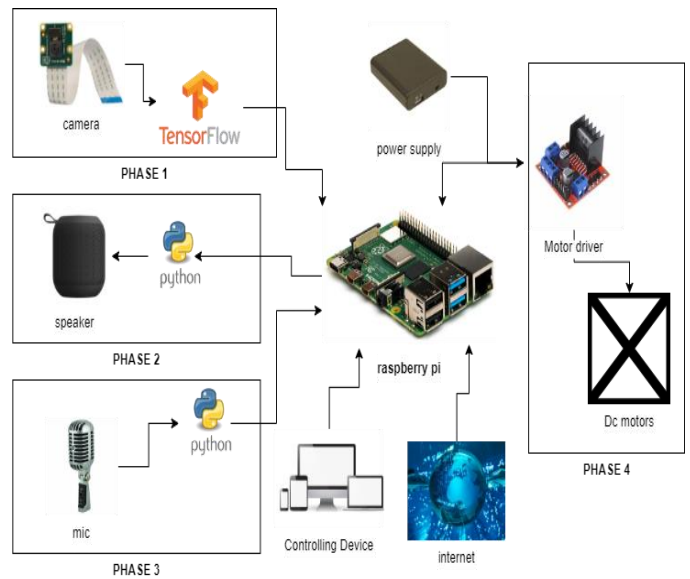

Fig 4: System Architecture

Working with Phases:

Phase 1: In this phase take input from the pi camera. The image is later processed with the help of OpenCV. We can't use the captured image as input. The image is processed and the processed image is used as input. The processing of an image takes place in raspberry pi. With the help of OpenCV, we use the EAST (Efficient and Accurate Scene Text) algorithm to process an image and obtain the desired input.

OpenCV's EAST text detector may be a deep learning model, supported a completely unique design and coaching pattern. EAST is defined as An Efficient and Accurate Scene Text Detector. With the help of the East, we detect text in an image and will use that text for further processing and to generate output. This Robo can be used for security purposes by detecting the face of the person standing at the door. It uses OpenCV to detect and identify the faces and assist.

Phase 2: This phase is the output phase of our project. Robo will give an output invoice. It will basically convert output text to voice. With the help of the speaker, it will give output to the user and it can be used for physically challenged people. This Robo uses our algorithm to give output. The optical character recognition (OCR) is the process that converts the scan or printed text images into the text format for further processing. This paper has bestowed the straightforward approach for text extraction and its conversion into speech. OCR algorithmic rule the mechanical or electronic conversion of pictures of typewritten, written or written text into machine-encoded text, whether or not from a scanned document, a photograph of a document, a scene-photo or from subtitle text superimposed on a picture.

Phase 3: This phase can be an optional phase for the project is about controlling the device remotely from different locations or areas by using a laptop or smartphone. This phase achieves the project goal flexibility and availability. This way it can be used from a remote location. For this, we just need a constant connection between the robot and the remote host and security password. You need to have a security password to access and give the command. In Raspberry, we need to install the Tight VNC server. The first time when we start it, it will ask for an 8-digit password as an access key to the security key. On the Windows side, we install Tight VNC and after that, we need the IP address and access key and we can command pi remotely.

Phase 4: This is again an output phase for the project, the robotic movement and behavior of the various kind of input response into action are provided here by the movement. Object detection and classification area unit major challenges for robotic modules. Recent years have provided good progress in object detection primarily because of machine learning strategies that became sensible and economical.
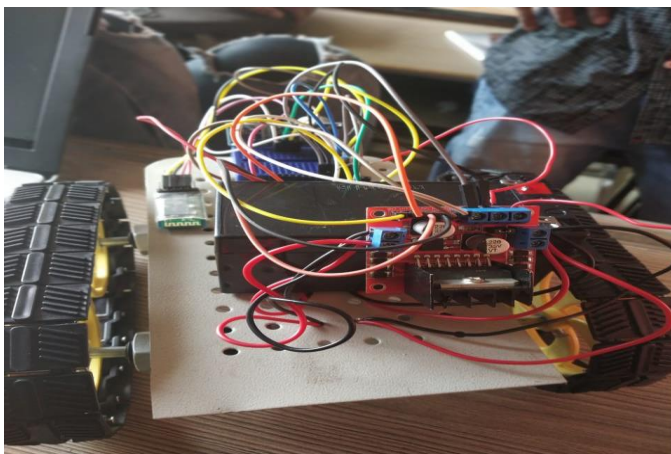
V. EXPERIMENTAL RESULTS

Fig 4.1: Back View

Fig 4.2: Side View

Fig 4.3: Side View

The above figure is the back view and side view of our mower. It shows all the connections which consist of a battery, Motor driver, Raspberry pi board, etc. are used in this project. This project has hardware connection with the help of wires this hardware plays an important role for working of this project can perform almost an accurately with the raspberry pi the algorithm which is used in this may make this project lag in facial recognition or the object recognition because of the frame rate that is been used in the raspberry pi system.

The valuable output for the Voice command is used in this project is for movement i.e. Forward, Back, Left, Right this is the basic movement command, the alternative features of this project is working as an assistant to a user it has the following command for accessing music, files, news, and some features of searching over the internet and getting the summarized content of the query that are been asked by the user. It can do the basic mathematical problem that is subtraction, multiplication, addition, division, as well as the square root and many features are been embedded in the system.

This system can work on minimum power supply of 5V that is need for the working of the raspberry pi system the raspberry pi board is the main circuit board of the system it controls all the divides, as well as the hardware that is attached to the system the mic works as the input source through which the input is passed to the system and through the raspberry pi the output, are been passed but the speaker as well as with the motor driver for the movement of the mower as per the direction that is commanded by the user.

Table 1: Voice Commands

Voice command	Movement
Forward	robot's body moves forward
Backward	robot's body moves backward
Left	robot's body moves left
Right	robot's body moves right

VI. CONCLUSION AND FUTURE WORK

The robot developed in our project are able to move in any direction like the front, back, left, right according to the voice command received from the user through mic as part of our hardware in this project. There is an autonomous voice command which can instantly make the robot move automatically without hitting any obstacle using the ultrasonic sensor. Our device will help users to give a uniform look to your lawn with ease. As well as it can also be used for the blind or physically challenged people by embedding this system into the wheelchair that will make an autonomous wheelchair.

REFERENCES

[1] Anurag Mishra, Pooja Makula, Akshay Kumar, Krit karan, and V.K. Mittal, "A voice-controlled personal assistant robot", Published in 2015 International Conference on Industrial Instrumentation and Control (ICIC).

[2] Xianghua Fan, Fuyou Zhang, Haixia Wang, and Xiao Lu, "The system of face detection based on OpenCV", Published in 2012 24ᵗʰ Chinese Control and Decision Conference (CCDC).

[3] Dyah Ayu Anggreini Tuasikal, Hanif Fakhrurroja, Carmadi Machbub "Voice Activation Using Speaker Recognition for Controlling Humanoid Robot", Published in 2018 IEEE 8th International Conference on System Engineering and Technology (ICSET).

[4] U Bharath Sai, K Sivanagamani, B Satish, UG Students, "Voice-controlled Humanoid Robot with artificial vision", Published in 2017 International Conference on Trends in Electronics and Informatics (ICEI).

[5] Renuka Kajale, Soubhik Das and Paritosh Medhekar, "Supervised machine learning in intelligent character recognition of handwritten and printed nameplate", Published in 2017 International Conference on Advances in Computing, Communication, and Control (ICAC3).

[6] X. Chen, A. L. Yuille, "A Time-Effect Cascade for Real-Time Object Detection: With applications for the visually impaired", 1st Intern. Workshop on Comp. Vision Applications for the visually CVACVI, June 20, 2005.

[7] P. Viola and M. Jones, "Robust Real-Time Face Detection", International Journal of Computer Vision, Vol. 57(2), pp. 137–154,2004

[8] https://www.raspberrypi.org/documentation/usage/gpio/