Department of Robotics and Automation

Jain Global Campus, Kanakapura Taluk - 562112

Ramanagara District, Karnataka, India

2021-22

A Report on

"Smart Assistant Bot "

Submitted by:

Aravind Vishwanath Chindi (19BTLRA002)

Manish Kumar Gupta (19BTRRA026)

Ashim Dongol (19BTRRA023)

Nwar Abo Daka (19BTRRA027)

Under the guidance of:

Dr. Kamala N

Department of Robotics and Automation

Faculty of Engineering and Technology

Jain Deemed to-be University

Abstract

In this project, we are going to build a mobile smart robot with smart features. The features include face recognition, voice recognition along with voice command, and object detection. The bot will be following the specific defined instructions. It will also detect the user face and related functionalities like greetings, etc

Along with face recognition, the robot will follow the user given commands for movement like come forward, go back, etc. In the project, we will be using ESP 32 as our main controller along with a camera module, microphone, etc. For the movement, there will be an L293D motor driver module for controlling motors.

In future, we will be adding up more features like navigation and path planning, iot interface, user emotion detection.

1. INTRODUCTION

Gesture recognition is a subject in technological know-how and language generation with the aim of interpreting human gestures through mathematical algorithms. Gestures can originate from any physical movement or country however usually originate from the face or hand. Users can use easy gestures to govern or have interaction with gadgets without bodily touching them. Many methods had been made using cameras and computer imaginative and prescient algorithms to interpret signal language. However, the identity and popularity of posture, gait, proxemics, and human behaviors is likewise the difficulty of gesture popularity techniques. Gesture popularity may be visible as a manner for computer systems to start to understand human frame language, hence building a richer bridge among machines and humans than primitive textual content user interfaces or maybe GUIs (graphical user interfaces), which nevertheless restriction the bulk of enter to keyboard and mouse and have interaction evidently with none mechanical gadgets. Using the idea of gesture recognition, it is viable to factor a finger at this factor to circulate accordingly. This may want to make conventional enter on gadgets such or even redundant.

A gesture is a movement that must be visible by a person else and has to carry a few pieces of information. Gesture is commonly taken into consideration as a motion of a part of the body, esp. a hand or the head, to express a concept or meaning.

In the early years the most effective manner to talk with a robot became software which required great difficult work. With the improvement in technology and robotics, gestures primarily based totally on reputation got here into life. Gestures originate from any physical movement or nation however normally originate from the face or hand. Gesture reputation may be taken into consideration as a manner for a laptop to recognize human frame language. This has minimized the want for textual content interfaces and GUIs (Graphical User Interface).

2. Objective

The integration of more and more functionality into the human machine interface of vehicles increases the complexity of device handling. Thus optimal use of different human sensory channels or gestures is an approach to simplify the interaction with in-car devices. Using this idea, a car-robot can be implemented whose navigation can be done wirelessly with the help of a Raspberry Pi.

This is the Autonomous primarily based totally project , wherein we specially make use of the Raspberry Pi, USB web camera and DC motor with Robot chassis to construct this Robotic vehicle setup. It has an internet digital digicam installed over it, through which we are able to get a video feed and the interesting component right here is that we are able to manage and move this robot from an internet browser over the internet. As it is able to be controlled the usage of web page, manner it is able to additionally be managed by the usage of the opposite smart devices wherein we are able to manage through the web page

The webcam will seize information with regards to its environment after which it becomes a preferred tool through the internet. The person can be observing this information at the screen on the person's end. According to the preferred movement, the person will manage the robot car through the internet web page available on the person's end.

3. LITERATURE REVIEW

Gesture recognition technologies are lots younger today. At this time there's lots of energetic studies withinside the subject and little withinside the manner of publicly to be had implementations. Several approaches were developed for sensing gestures and controlling robots. Glove's primarily based totally method is a well-known method of recognizing hand gestures. It utilizes a sensor connected to a glove that immediately measures hand movements.

A Gesture Controlled robot is a form of robot which may be controlled by hand gestures and now no longer the old style manner by the use of buttons. The consumer simply wishes to put on a small transmitting tool on his hand which incorporates a sensor which is an accelerometer in our case. Movement of the hand in a particular course will transmit a command to the robot so one can then circulate in a particular course. The transmitting tool consists of a Comparator IC for assigning right tiers to the input voltages from the accelerometer and an Encoder IC which is used to encode the four bit information after which it is going to be transmitted through an RF Transmitter module.

At the receiving end an RF Receiver module will receive the encoded records and decode it with the aid of the usage of a decoder IC. This information is then processed with the aid of using a microcontroller and surpassed onto a motor driver to rotate the automobiles in a unique configuration to make the robot pass withinside the equal course as that of the hand.

A hand-gesture-primarily based totally manipulated interface was delivered for navigating a car-robotic in [1]. A 3-axis accelerometer is followed to record a user's hand trajectories. The trajectory record is transmitted wirelessly through an RF module to a computer. The obtained trajectories are then categorized to certainly be considered one among six manipulate instructions for navigating a car-robot. The classifier adopts the dynamic time warping (DTW) set of rules to categorize hand trajectories. Simulation outcomes display that the classifier should achieve 92.2% accurate rate.

A novel, non-contact, pointing interface is being evolved for manipulation of non-protection important systems inside an automobile with the goals of enhancing protection, reducing manufacturing fee and enhancing the benefit of driver migration among unique automobiles in

[2]. A driver operates the interface thru an onscreen cursor the use of pointing gestures to be diagnosed through a computer vision device. This paper describes the imaginative and prescient subsystem accountable for detection and monitoring of the driving force's arms. To be strong, it should come across and track beneath various lighting fixtures situations without any previous assumptions regarding the color of the arms or clothing. Adaptive foreground and history fashions are used for segmentation and a strong geometrical hand version is hired for monitoring. The gadget is validated operating at speeds near real-time on a preferred PC the use of image sequences captured inside a car.

The integration of increasingly more capability into the human device interface of cars will increase the complexity of device dealing with in [3]. Thus the most excellent use of various human sensory channels is a technique to simplify the interplay with in-automobile gadgets. This manner of personal convenience will increase as much as distraction might also additionally decrease. In this paper a video primarily based totally actual time hand gesture popularity device for in-automobile use is presented. It evolved in the course of significant usability studies. In aggregate with a gesture optimized Human device interface it allows intuitive and powerful operation of a variety of in-automobile multimedia and infotainment devices with hand poses and dynamic hand gestures.

Envision to feature context attention and ambient intelligence to edutainment and laptop gaming packages became widespread in [4]. This calls for mixed-truth setups and ever-better stages of immersive human-laptop interaction. Here, the focal point is the computerized popularity of natural human hand gestures recorded with the aid of using inexpensive, wearable movement sensors.

To look at the feasibility of this approach, an educational parking recreation was selected with 3-d snapshots that employs movement sensors and hand gestures as its sole game controls. The implementation prototype is primarily based totally on Java-three-D for the snapshots show and at the CRN Toolbox for sensor integration. It indicates very promising consequences in exercise concerning recreation appeal, participant satisfaction, extensibility, ease of interfacing to the sensors, and – remaining however now no longer least – sufficient accuracy of the real-time

gesture popularity to permit for easy recreation control. An preliminary quantitative overall performance assessment confirms those notions and provides in addition assistance for the setup.

The number one and secondary riding challenge collectively with Human Machine Interface developments and problems which can be driving car consumer interface designers to take into account hand gesture recognition as a sensible opportunity for consumer controls are defined in [5]. A wide variety of hand gesture recognition technology and programs for Human Vehicle Interaction also are mentioned consisting of a precis of modern car hand gesture reputation research.

Hand gesture recognition is an important way for Human-Robot Interaction. Sign language is the maximum intuitive and direct manner to communicate for impaired or disabled humans. Furthermore, emotional interplay with people is appropriate for robots. In this paper [6], hand gesture recognition and emotional reputation of an incorporated device might be defined which has the potential to track more than one human at the same time, to apprehend their facial expressions, and to perceive social atmosphere. Consequently, robots can without difficulty recognize hand gestures and facial features with emotion versions of various humans, and might respond properly. A combining hand gesture reputation set of rules which mixes wonderful recognizers has been studied. These recognizers together determine the hand's gesture through a system referred to as combinatorial approach recognizer (CAR) equation.

Automobiles have become more and more critical in each day of life. However, human beings commonly want to feed masses of effort and time to get their required services. On one hand, a car itself includes many bodily methods to gain its conventional functionality; on the other hand, its miles are being included with increasingly more sensors and actuators, which makes it a regular cyber physical system. In order to offer pleasant and human-centric offerings for users, the Intelligent Cyber-Physical System for automobiles (iCPS-Car) for vehicles was proposed in [7]. iCPSCar integrates human beings, automobiles and cyber areas collectively and offers herbal interplay manners, and personalized and continuous services.

It attempted to narrow down the gap between the real world and synthetic environment in [8]. For that purpose, an immersive driving car simulation was developed that combines tangible

tools with a mixed reality environment. As a tangible tool, a real physical mini car was deployed that employed an arduino sensor inside. To provide an immersive sensation among users and arduino mini cars, a camera belonging to a smart phone was employed in front of our car that will create a real driving-sensation since the user will feel like he is sitting and driving inside the car. As a car controller, a natural user interface (NUI) controller was implemented that employed the widely used RGBD sensor, Kinect.

Multiple devices while driving steals drivers' attention from the road and is becoming the cause of accidents in 1 out of 3 cases. Many research efforts are being dedicated to design, manufacture and test Human-Machine Interfaces that allow operating car devices without distracting the drivers' attention. A complete system for controlling the infotainment equipment through hand gestures is explained in this paper. The system works with a visible-infrared camera mounted on the ceiling of the car and pointing to the shift-stick Area, and is based on a combination of some new and some well-known computer vision algorithms in [9]. The system has been tested by 23 volunteers on a car simulator and a real vehicle and the results show that the users slightly prefer this system to an equivalent one based on a touch-screen interface.

Humans and machines do not interface well. In an attempt to bridge the gap between humans and the systems they interact with, a plethora of input methods have been devised: keyboards, mouse, joysticks, game controllers and touch screens are just a few examples. Unfortunately, none of these devices remove the barrier between man and machine. With the Magic Glove control system in [10], the aim is to remove this obstruction by allowing the user to control a hardware device using natural gestures. The Magic Glove takes advantage of a multitude of sensors to capture hand movements and uses this information to control a device – in this case, a modified RC car. The goal of this paper is to capture simple hand gestures from the Magic Glove and use that input to wirelessly control a modified RC car. Controlled variables include speed, steering, lights and sounds using a combination of flex, force and gyroscopic sensors. Multiple variables are controlled simultaneously as Magic Glove outputs a constant control signal.

This paper [11] presents a method for two-hand pose recognition based on skeleton information, aiming at the problem of low recognition rate and poor robustness in the field of human computer interaction by single hand. This method consists of two steps: two-hand positional

information extraction and gesture recognition. In the first step, we utilize the Kinect depth image to acquire the position of both hands. The second step is the highlight of the proposed method, it locates the palms by hand nodes, extracts the right hand movement information which is trained by a Hidden Markov Model. This method has been verified by an experimental car control system and demonstrated good robustness in a complex background environment.

To ensure safety and usability of advanced in-car cockpit solutions, prospective evaluation during early prototyping stages is important, especially when developing innovative human cockpit-interactions. In this context, highly realistic test environments will help to provide reliable and valid findings. Nevertheless, real car driving studies are difficult to control, manipulate, replicate and standardize. They are also more time consuming and expensive. One economizing suggestion is the implementation of immersive driving environments within simulator studies to provide users with a more realistic awareness of the situation in [12]. This paper discusses research investigating the influence of immersive driving environments. Three interaction modalities (touch, spin controller, free-hand gestures) and two levels of immersivity (low, high) are examined to examine this methodology.

The recently developed Kinect sensor has opened a new horizon to Human-Computer Interface (HCI) and its native connection with Microsoft's product line of Xbox 360 and Xbox One video game consoles makes completely hands-free control in the next generation of gaming in [13]. Games that require a lot of degrees of freedoms, especially the driving control of a car in racing games, are best suitable to be driven by gestures, as the use of simple buttons does not scale to the increased number of assistive, comfort, and infotainment functions. In this paper, Mamdani type-I fuzzy inference system based data processing module is proposed which effectively takes into account the dependence of actual steering angle with the distance of two palm positions and angle generated with respect to the sagittal plane. The FIS output variable controls the duration of a virtual "key-pressed" event which mocks the users pressing of actual keys assigned to control car direction in the original game. The acceleration and brake (deceleration) of the vehicle is controlled using the relative displacement of left and right feet.

The notion of developing thought controlled devices (games, robots, cars etc.) is becoming increasingly popular with the introduction of low cost commercial headsets that record neuro

electric activity and the extensive research in the area of Brain Computer Interfaces (BCIs). In this paper [14], we study the feasibility of using a commercial low cost EEG amplifier which has only a limited number of electrodes, to develop a motor control BCI system. The objective is to extract brain activity responsible for direction specific imagined and executed motor activity, which can be used to identify the motor task performed by the user using the simultaneously recorded EEG.

Recently it was issued that the interaction technology for driver's gesture recognition in vehicular environment. Drivers want to control the multimedia system, air conditioning system and other applications which are equipped in head units on the dashboard through simple hand motion. But, limited under the safety driving condition, gesture cognition while being in the car has many problems due to specific conditions on the road. In a car, according to vehicle status such as forward and backward moving or horizontal tilting, data errors of the motion sensors are caused and unexpected driver's motion will be registered. In this paper [15], it is proposed that the system model for the gesture interaction between users and vehicle are defined and the data processing process including specific hardware structure is adopted in order to reduce the motion sensing errors. The past few years has shown a sudden spurt in the field of human computer interaction. The days of using a mouse to control a computer is almost obsolete and people now prefer to use touch screens and more recently, air gestures, for the same. However, the use of gestures is not limited to computers alone. It finds its application in controlling televisions and other home appliances as well. This paper [16] explores one such application where in air gestures could be used to control automobiles. The paper describes a novel method to not only give directions but also password-protect and use special features of the vehicle, using gestures. The complete algorithm was developed using video processing on MATLAB 2011b and was found to be 3.6 times faster than its predecessor algorithms. The same was tested in real-time as well, using a robot prototype and satisfactory results were obtained.

In this contribution [17], a novel approach to transform data is presented from time-of-flight (ToF) sensors to be interpretable by Convolutional Neural Networks (CNNs). As ToF data tends to be overly noisy depending on various factors such as illumination, reflection coefficient and distance, the need for a robust algorithmic approach becomes evident. By spanning a three

dimensional grid of fixed size around each point cloud we are able to transform three dimensional input to become processable by CNNs. This simple and effective neighborhood preserving methodology demonstrates that CNNs are indeed able to extract the relevant information and learn a set of filters, enabling them to differentiate a complex set of ten different gestures obtained from 20 different individuals and containing 600.000 samples overall.

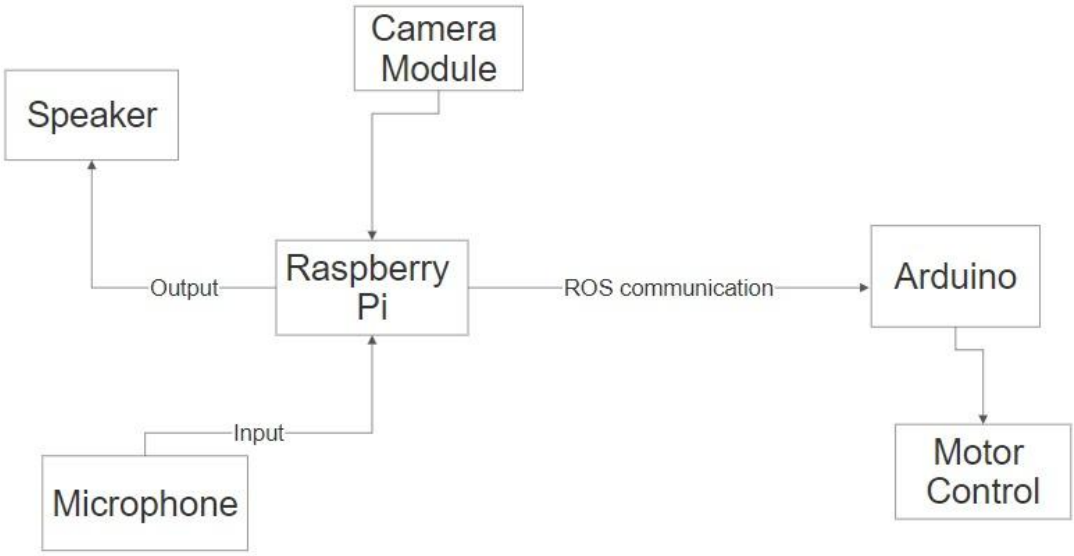
In this paper [18], we introduce a hand-gesture-based control interface for navigating a car-robot. A 3-axis accelerometer is adopted to record a user's hand trajectories. The trajectory data is transmitted wirelessly via an RF module to a computer. The received trajectories are then classified to one of six control commands for navigating a car-robot. The classifier adopts the dynamic time warping (DTW) algorithm to classify hand trajectories. Simulation results show that the classifier could achieve 92.2% correct rate.

A research platform has been designed for a perceptually guided robot, which also serves as a demonstrator for a coming generation of service robots in [19]. In order to operate semi autonomously, these require a capacity for learning about their environment and tasks, and will have to interact directly with their human operators. Thus, they must be supplied with skills in the fields of human-computer interaction, vision, and manipulation. GripSee is able to autonomously grasp and manipulate objects on a table in front of it. The choice of object, the grip to be used, and the desired final position are indicated by an operator using hand gestures.

Gesture interfaces are gaining relevance for human-machine communication, since it is expected that they make interaction more intuitive. Particularly vision based approaches are widely preferred. This paper [20] describes a novel vision based real-time gesture recognition system, designed for operating in an automotive environment. It is used within an application for retrieving traffic news and emails from a message storage. Image processing and pattern matching techniques, specially adapted to the complex environmental conditions, represent the system's basics.

Proposed System:

The block diagram of the system is illustrated in the given figure.

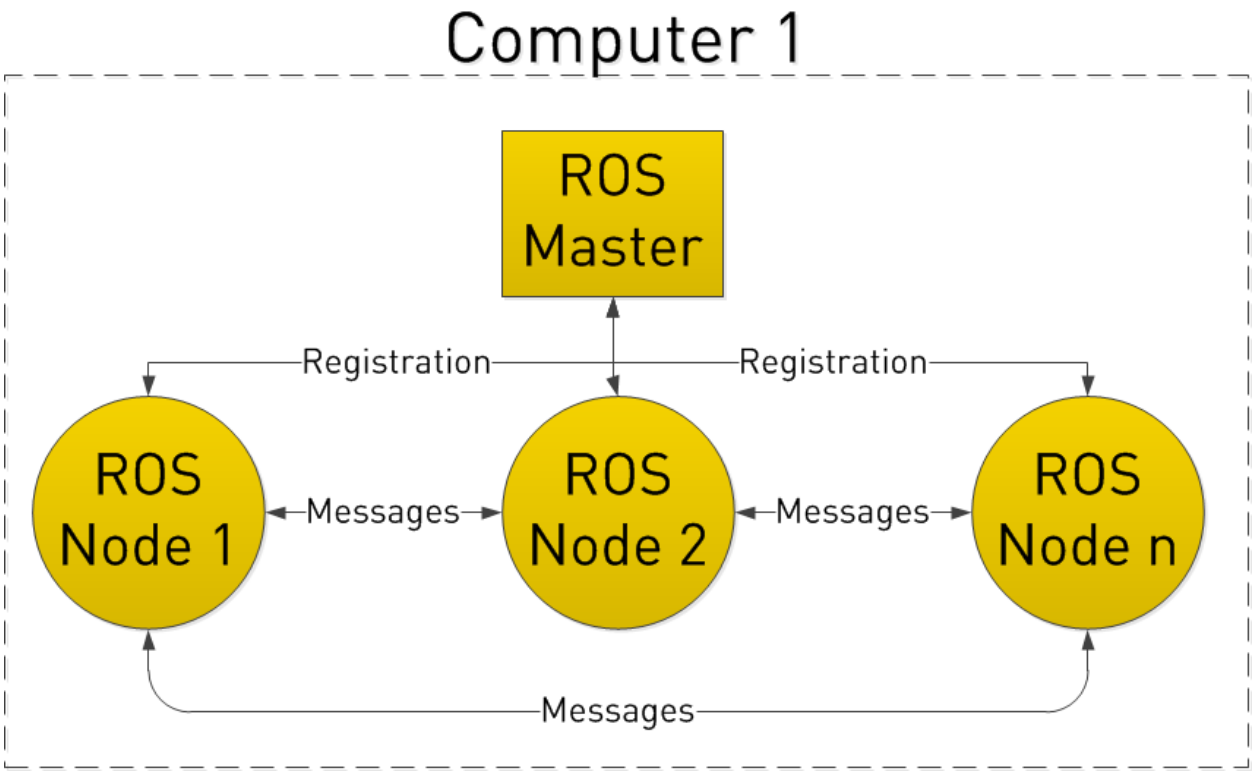


A Raspberry Pi board, an Arduino board, a speaker, a microphone, motor drives, and a camera module compose up the system. The camera module receives and processes the video signal-like gesture. If a specific action is to be done after processing the video feed, the raspberry pi instructs the arduino to move the motor as instructed. The Raspberry Pi and Arduino communicate via ROS, with the Raspberry Pi serving as the master and the Arduino serving as the slave.To conduct particular moves, the Raspberry Pi can also receive voice input from the microphone.

Software used:

a. ROS:

The Robot Operating System (ROS) is a collection of software libraries and tools that assist in the development of robot applications. From drivers to cutting-edge algorithms, with robust developer tools. ROS + Raspberry Pi is an incredibly strong combination for building smart robots using a relatively low-cost and compact electronic board embedded in the robot.



b. Python:

Python was the programming language deployed on the Raspberry Pi for this project.Python is a high-level, interpreted programming language for general-purpose applications. Python has a design philosophy that prioritizes code readability, which includes a lot of whitespace. It has structures that allow for clear programming at both small and large sizes.

Python has a dynamic type system and memory management that is automatic. Object-oriented, imperative, functional, and procedural programming paradigms are all supported. It also includes a huge standard library.

c. Arduino Ide:

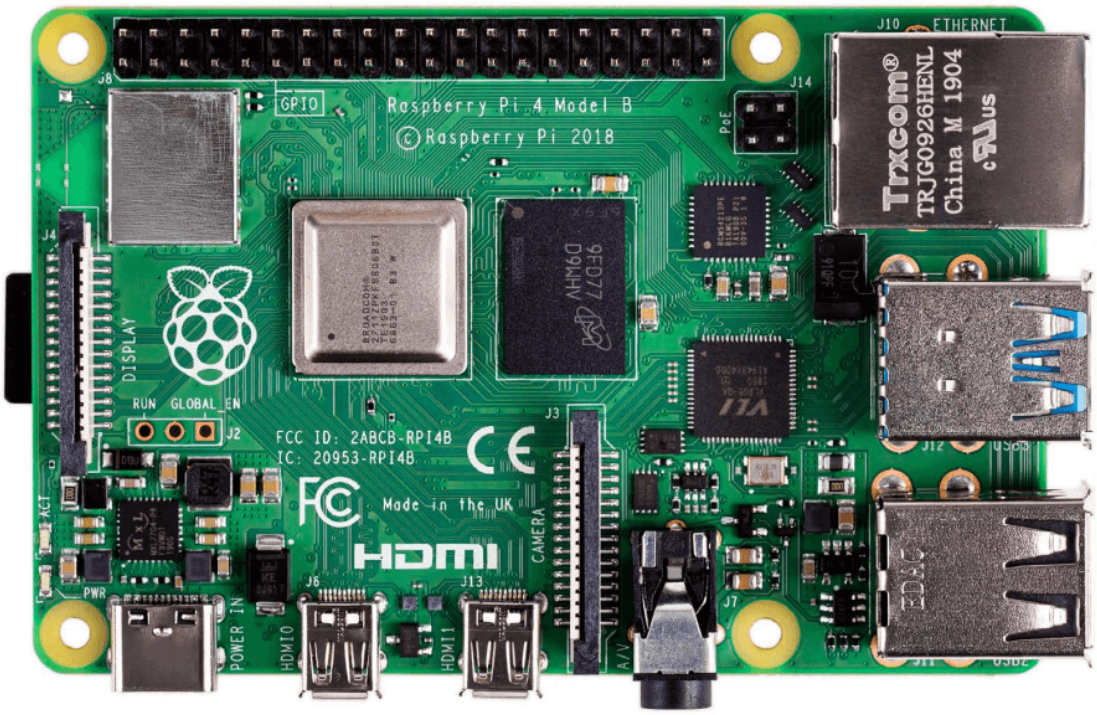
The Arduino board is programmed using the Arduino IDE.The Arduino Integrated Development Environment (IDE) is a C and C++-based bridge system for Windows, macOS, and Linux. This is used to write and upload computer programmes to Embedded system boards and other entrepreneurship development boards that enable third-party cores. We may use this IDE to programme the arduino board to control the motion of the robot's wheels based on the command.

Hardware used:

a. Raspberry pi:

The Raspberry Pi Foundation, a UK nonprofit that strives to educate people in computing and make computing education more accessible, has created a series of single-board computers known as the Raspberry Pi.The Raspberry Pi is a low-cost computer the size of a credit card that connects to a computer monitor or television and utilizes a conventional keyboard and mouse. It's a capable small device that allows individuals of all ages to learn about computers and programming languages like Scratch and Python. It also aids in the development of hardware projects.

Raspberry Pi is used to communicate between the ROS programmes and it also helps to control the Arduino Board.

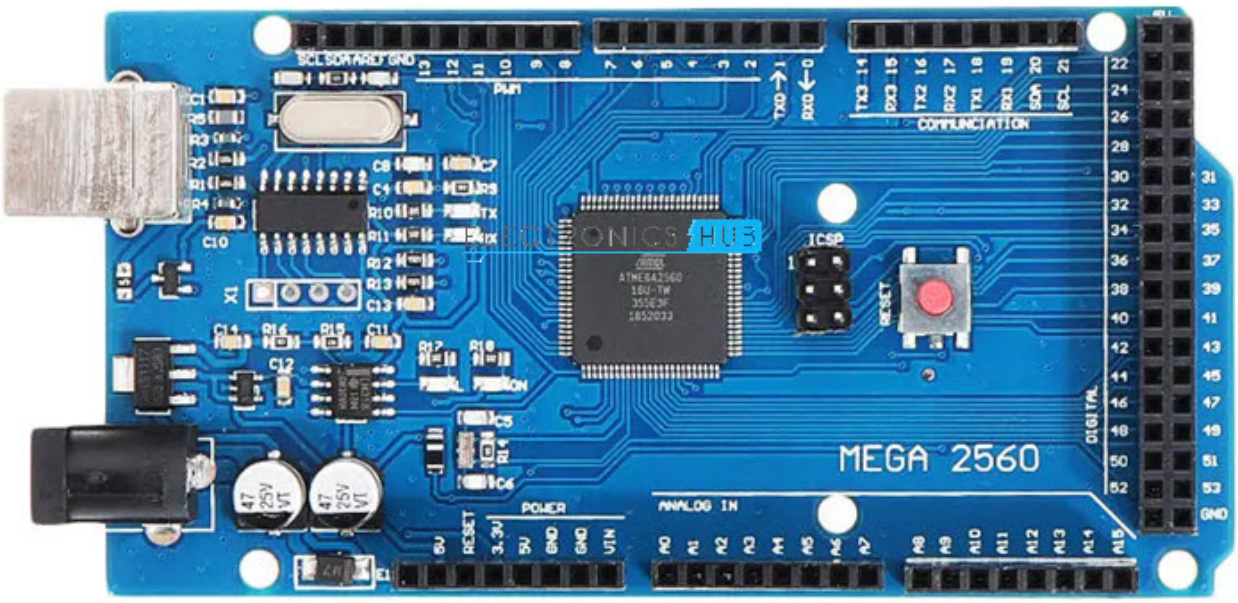


Raspberry Pi communicates with Arduino through the Serial Communication node created by ROS. rosserial is a protocol for wrapping standard ROS serialized messages and multiplexing multiple topics and services over a character device such as a serial port or network socket. RosSerial provides a ROS communication protocol that works over your Arduino's UART. It allows your Arduino to be a full fledged ROS node which can directly publish and subscribe to ROS messages, publish TF transforms, and get the ROS system time.

b. Arduino :

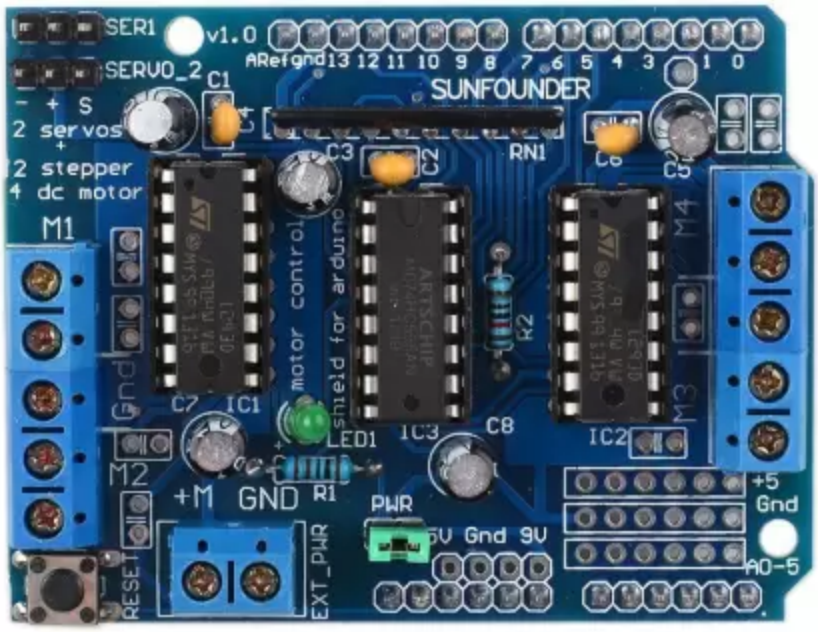
Arduino is an open platform for creating electronics projects that is simple to use. Arduino boards are essential for a variety of projects. Arduino is a commonly used board for small hardware projects. It is a very easy to use board containing several analog and digital pins for interfacing other devices or sensors to carry a certain operation.

The motor function of the project is controlled by an Arduino. This arduino board is connected to four dc motors. A Raspberry Pi board is also linked to the board. The Raspberry Pi serves as the master, while the Arduino serves as the slave.



+c. L293D Motor Driver IC :

The DC motor's spinning direction can be controlled by changing the polarity of its input voltage. A common technique for doing this is to use an H-Bridge. An H-Bridge circuit contains four switches with the motor at the center forming an H-like arrangement.



The full-featured L293D motor driver shield can control up to four bi-directional DC motors with 8-bit speed selection, two stepper motors, and two servo motors. The power supply to the driver IC can be used for both the shield and Arduino or the two can use separate power supplies. If sharing a common power supply, a power jumper must be placed on the driver IC. The power supply can be provided through Arduino's USB port, its DC jack, or from the driver IC's 2-pin EXT-PWR block.

To connect the DC motors, two five-pin screw terminals are at the edges of the shield. These terminals are labeled as M1, M2, M3, and M4. Those DC motors with voltage ratings of between 4.5 to 25V can be connected to these terminals and will deliver power up to 600 mA.

d. Raspberry Pi IR-CUT Night Vision Camera :

This is the camera that was used for the project. It also enables night vision. It can be utilized in low-light conditions and still identify and recognise faces.The IR LEDs are powered directly from the CSI port, and are capable of lighting an area at a distance of up to 8m! In testing, the best images were captured at a distance of 3m to 5m. The camera also features an adjustable 3.6mm focal length and 75.7-degree viewing angle.

This Raspberry Pi night vision camera uses the same OV5647 as the standard Raspberry Pi camera, and is, therefore, able to deliver a crystal clear 5MP resolution image or 1080p HD video recording at 30fps!

e. Battery :

It provides the voltage required for the Raspberry Pi and the Arduino board to function properly. We used a 12 volt battery, which is higher than the needed voltage, thus the voltage is decreased before powering the board to avoid a short circuit.

f. DC motors:

These are the motors that rotate the robot's wheels. Motor drivers attached to the Arduino board are used to drive these motors.

g. Ultrasonic Sensor:

This sensor is used to determine the robot's distance from an impediment. It is used in this project to measure the distance between obstacles and to detect the edge of a table or other surfaces. It could be stairwells, dining tables, or study tables.

h. Servos:

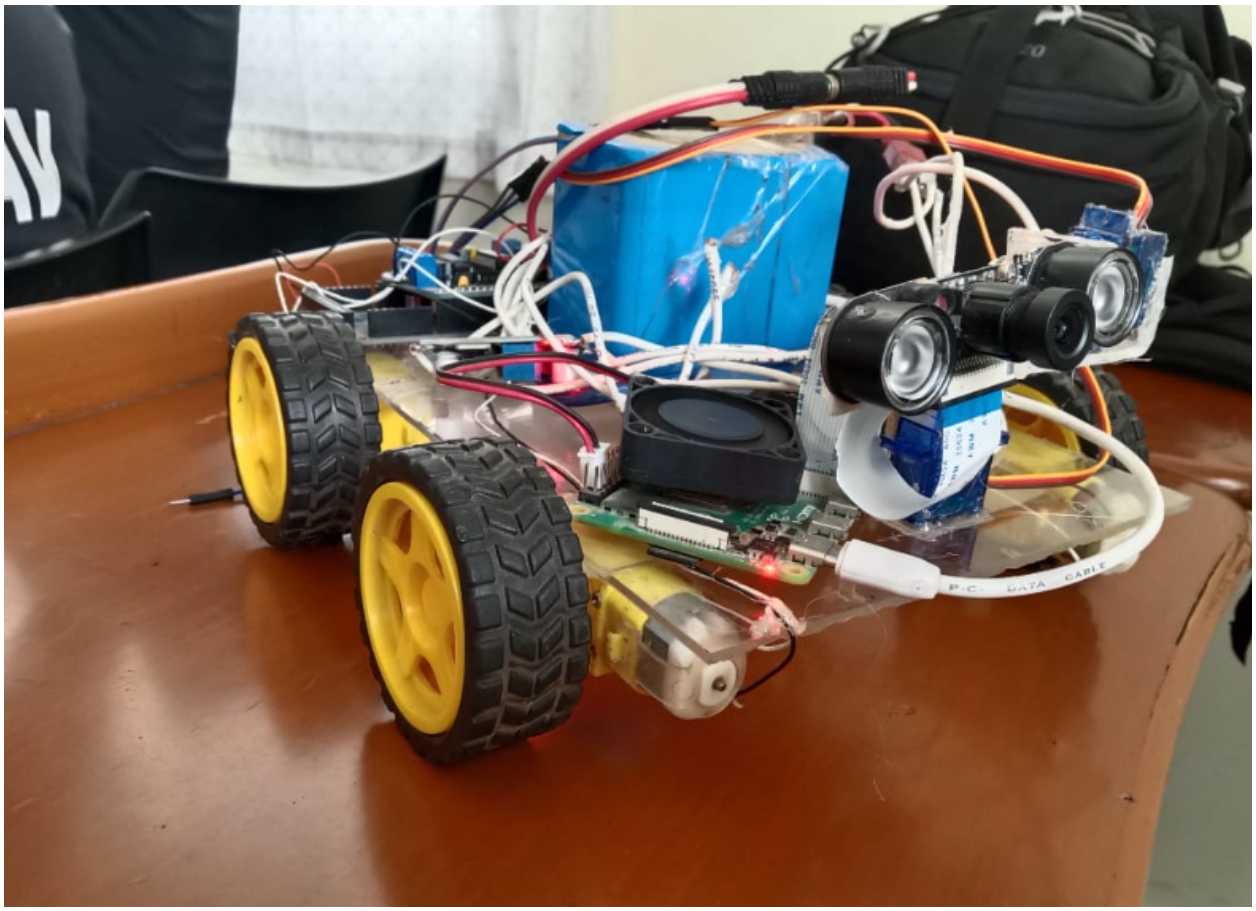
These are the motors that regulate the camera movement of the robot. To recognise and search for faces and bring them into the robot's frame, the camera must be oriented such that the entire face appears on the screen. Two servos are involved in this action, one for horizontal rotation and the other for vertical rotation.

i. Microphone:

The user's vocal commands are received through a microphone. It's linked with the Raspberry Pi. When the Raspberry Pi receives voice, it processes it and looks for equivalent commands if they exist. The command is then delivered to the Arduino board, which performs the desired activities.

5. Results and Discussion:

After connecting all the board and motors with a plane board the robot looked like:



6. Conclusion:

The gesture system was created to show that intuitive, easy gesture-based input could be applied to a wide range of applications, including hardware and software.The car makes the correct movements for the pre-programmed and calibrated hand gestures.It is equipped with a camera that may be used for surveillance, facial recognition, and hand gesture detection. The gesture-controlled surveillance vehicle has the ability to begin to close the gap between human concepts and machine reactions.The Raspberry Pi receives data from the camera's recording of hand movements. The Raspberry Pi identifies the gesture input and creates an equivalent command signal. The values are subsequently sent to the Arduino board, which ultimately sends them to the L293D motor driver. As a result, the data collected from the motor driver is used to control the motors.

In this speech control experiment, the car is also controlled by voice. A robot car is a mobile robot that can be controlled using voice instructions.Voice commands such as 'Forward,' 'Stop,' 'Left,' 'Right,' and 'Backward,' among others, will be recognised by the Raspberry Pi's speech recognition programme. Voice commands are used to control the robotic car through the microphone.

7. Future Scope:

The surveillance robot can be used for a variety of purposes. Its small size allows it to blend in with its surroundings, and it may be used to glance around and retrieve data from the environment.There are countless changes that can be made to the current design and technology, as well as numerous new features that can be introduced. Temperature sensor, pressure sensor, heat sensor, position sensor, and proximity sensor are all examples of sensors that can be used to use robots in many fields. A wireless network can be used to create a multipurpose robot with applications ranging from surveillance and home security to industrial applications where the user does not need to be physically present at the work site but can do it from his home.A

thermal camera can detect energy produced by bodies, making it helpful for defense equipment such as detecting attackers on the ground.