Voice Controlled Autonomous Rover Interfaced with Robotic Arm having Soft Gripper

Dr. E. Logashanmugam[1], K. Joseph Pranadeer Reddy[2] and J. Mohan Satish[2]

1Professor at Dept. of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, INDIA
2 Student at Dept. of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, INDIA

***ABSTRACT*— Robotics has become a booming area of research in the field of Engineering. It holds great potential in reducing human effort by performing tasks faster while maintaining operational accuracy. The Robotic arm has shown a prominent role in the field of robotics to aid as an artificial human hand using its manipulator and gripping mechanism. Also, recent advances in the field of Artificial Intelligence has greatly helped to make the instruments intelligent. Our proposed project integrates these fields of technology which can be implemented in numerous sectors. This project comprises three major portions- ground rover, robotic arm, and image processing-based control. The rover has a rocker-bogie mechanism which makes it a robust vehicle to travel across various terrains. A robotic arm is attached to the rover allowing it to manipulate objects along its path. The functioning of the rover and the robotic arm is governed by the image processing mechanism, enabling the device to identify the object to be manipulated. To promote the ease of access while providing input to the model, instead of furnishing with conventional remote buttons, the model is interfaced with voice recognition capability to take voice commands as input from the user. Our project is designed to distinguish and categorize the objects specified by the user and bring the object to the user. The scope of application for this project is versatile, not limiting its functionality to a fixed application.**

Keywords*— Rover, Robotic Arm, Rocker-Bogie mechanism, Object detection, Raspberry Pi microprocessor

I. Introduction

The use of robots has made life easier. The development of human assistants has also been taking place at a rapid rate. To promote this development of human assistants at work, home, and elsewhere, we have developed a prototype of human assistant inspired by the exo-planetary rovers and 'SPOT'- Boston Dynamics robot. Building such systems means building a remote version of ourselves.

II. existing system

Very few human assistants have been made using rovers. A reasonable number of works have found in Literature. Although the techniques and equipment used in this project have already been proposed and are in use, we have used a unique combination of algorithms and mechanical structure which makes our proposed project novel.

Many Algorithms and works have been proposed and developed to achieve a better Robotic System. Nevertheless, every proposed system lacks a few features. Even though the drawbacks of one system have been compensated in the other, No system as a whole can work autonomously, reducing the effort put in by the user.

Most of the work done on object tracking robots make use of sensors such as infra-red sensors, photodiodes, and color sensors to detect the objects. The use of such sensors adds to the size of the model. Furthermore, the system becomes proportionately complex. Power handling becomes a frequent problem in such cases. Also, the robotic arm attached to the vehicle makes its movement unstable due to carrying the load on non-uniform terrains.

III. Overview of the proposed system

A. Proposal

The sole purpose of this project is to search for the object specified by the user's voice command, collect the object using the Robotic Arm, and bring it back to the user. This procedure involves a lot of complications. The audio signals from the voice command have to be converted to machine language so that the processor can understand the given command and start searching for it. The image processing technique used has to be more efficient, utilizing less computational power, so that there is no breakdown while performing. Once the object is detected, the rover has to be aligned in the direction of the object and then move towards it. Finally, identifying the coordinates of the object, the robotic arm has to grab the object and the rover has to bring it back to the user. Power supply to the entire system plays an essential role to achieve the goal.

- A Rover attached with a Robotic Arm.
- The Rover is equipped with the Rocker-Bogie mechanism to handle rugged terrains.
- Robotic Arm is used for handling various objects.
- Camera and Ultra Sonic Sensor help in environment analysis which aids Path planning and Object detection.
- The commands are given as voice input.
- The entire functioning is handled by Raspberry Pi microprocessor.

B. Novelty

- Our proposed project integrates all the requirements to compensate for the specified drawbacks.
- The rover's movement is made stable using the rocker-bogie mechanism.
- Instead of using multiple sensors, a camera is used for object detection.
- The system is designed to make decisions on its own, once the user specifies the object.

IV. Materials and Techniques

Different materials and techniques used in the projects have been elaborated below.

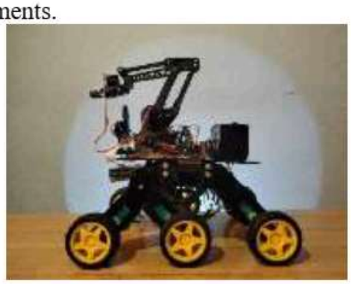
A. ROCKER BOGIE MECHANISM

This mechanism is used to handle one of the drawbacks which could not be addressed by many proposals mentioned in the literature survey.

The stability of the robot is hindered while working in a non-uniform terrain, especially when carrying any load. Even though the concept of the rocker-bogie system has been there for years and has been widely used in exo-planetary rovers, it hasn't found its wide usage in everyday house rovers.

A rover has to be agile and stable (because if it tips over, it can't get itself up). The six wheels on the rover act as two little tripods. The different pivots which allow three points of contact to be on the surface bearing the weight on three different points on either side. For building this prototype, rubber tires have been used, which can be replaced with other materials based on the type of surface on which it will be employed.

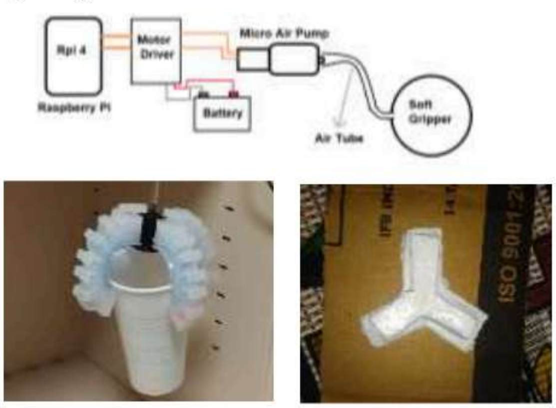
Hence, the use of this system provides great stability to the rover, making its application open to various environments.



B. ROBOTIC ARM

A 3D printed Robotic ARM has been used, which is driven by servo motors to attain a precise control over the Manipulation. It has 3 degrees of freedom about its base, shoulder, and elbow, thus mimicking a human hand, enhancing dexterity.

To achieve a better grip over the objects to be collected, a Soft Gripper can be used. Such grippers are made using "Eco flex 00-30 super soft platinum silicon rubber", which are driver by pumping air into the gripper. Hence, depending on the object, proportionate air can be pumped to grasp the objects. This mechanism plays a crucial role while handling fragile objects.



Soft Gripper

Soft pneumatic "finger" is made of flexible material. The grasping action is similar to the finger of the person. It is flexible and can automatically wrap the product, thereby avoiding physical damage to the product. Enabling grasping by a) actuation, b) controlled stiffness, and c) controlled adhesion. Challenges for soft grippers include miniaturization, robustness, speed, integration of sensing, and control.

To handle sharp iron pieces, an electro-magnetic end effector can be used, in order to improve efficiency. Such end effectors can be made by passing electricity to an iron plate fixed at the gripper point.

C. OBJECT DETECTION TECHNIQUE

Two object detection techniques have been used in this proposal.

Initially, the SURF feature extraction algorithm was used to identify objects. SURF employs a sliding window to perform object detection. Such a Deformable Parts Model uses a disjoint pipeline to extract features, classify regions and predict bounding boxes for high scoring regions, etc. This process can be made efficient by the usage of a single Convolution neural network, to replace these disparate parts by performing the tasks of feature extraction, bounding box prediction, etc., concurrently.

Thus, a Convolutional Neural Network using the TensorFlow library has been employed to achieve results with high accuracy. Instead of training a model from start, we used an existing model and fine-tuned it for our requirement.

Secondly, a Haar Cascade is a classifier used to detect the object for which it has been trained for, from the source. Haar Cascade is a machine learning object detection algorithm trained by superimposing the positive images over a set of negative images. This pre-trained network available in OpenCV enables the users to perform various detection tasks. Detecting things like faces, cars, smiles, eyes, and license plates for example are all pretty prevalent.

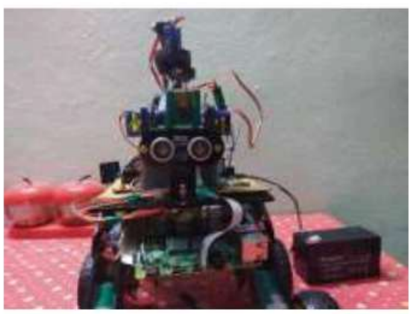
D. COMMUNICATION OF VOICE COMMANDS

An Android Application has been developed to communicate with the processor which controls the whole model. To avoid extra hardware and improve the user's comfort to access the model, an app is developed, instead of a physical remote controller. Usage of the app also allows the user to modify/develop the app quickly in accordance with the needs. The APP is designed to have 'Manual Control' over the Model, as well as give voice commands so that the system works autonomously. The voice commands are converted into strings by the app, which is passed to the processor. No special programming knowledge is needed for using this framework. The Logical construction has been designed carefully; bug-free.

Web Server has been used to establish a means of communication between Raspberry Pi and Mobile App. Bluetooth Communication can also be used but it has been unsteady and does not provide long-range communication. Different port IDs have been assigned for performing multiple functions. This long-range stable communication is initiated by Raspberry pi.

E. OTHER COMPONENTS

- CAMERA: It is an important component that governs the processing of visual information to the processor. As eyes are to a human, a Camera is to a system. All the image processing algorithms are dependent on the information from this component.
- ULTRASONIC SENSOR: Ultrasonic sensors measure distance by using ultrasonic waves. Ultrasonic Sensors measure the distance from the obstacle by measuring the time between the emission and reception of the wave.

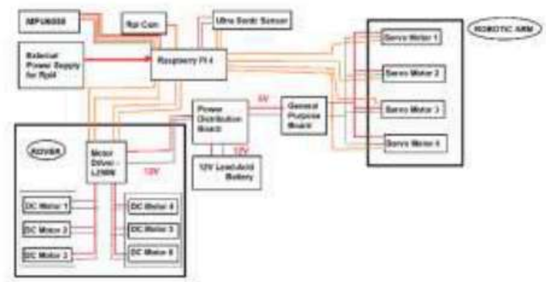


- IMU SENSOR: An Inertial Measurements Units sensor is used to understand the orientation of a body in three-dimensional space. As the rocker-bogie mechanism has been used to withstand elevations and depths on the surface, an MPU6050 has been used to analyze them along with the velocity.
- POWER SUPPLY: A 12V, 1.3Ah Lead-Acid External power supply is used to power up the components. The battery is charged using 'Battery Eliminator'. This battery on a full charge can run the Rover up to 3hrs, depending on the usage. A Power distribution board is used to regulate the voltage into 12V, 5V, and 3.3V. The distribution board has a flexible power input system. The board can be powered by using either wire connected to a battery or using an Adaptor compatible with the board. Raspberry Pi is powered using an external power source as this processor demands high power consumption.

V. CIRCUIT DESIGN

Raspberry pi 4 is connected to an external power bank. The battery is connected to the Power distribution board from where the power is split. The motor driver consumes 12V and gives 10V output to the DC Motors (using up 2V for Voltage Regulation and Heat Control is achieved using Aluminum heat sink attached to the Motor Driver. Servo Motors draw a Voltage of 5V.

The necessary control Signals are sent by Rpi4. The Servos are connected to PWM pins of Rpi4, thus controlling them. Ultra-sonic sensor draws power from the Raspberry pi power pins. Pi Camera doesn't need any special power. It stays active as long as Raspberry Pi gets the necessary Power. The web server is activated by using Raspberry Pi. Heat sinks have been placed on Raspberry Pi, to control the processing temperature. The Algorithms have also been designed to draw less computational Power and faster processing.



VI. COMPUTATION WORK

A. Rocker-Bogie Mechanism

The control over the wheels is obtained using an L298N Motor Driver. All the six wheels have been connected to this driver. Necessary soldering work is done to distribute the power supply to all the wheels. Using the Enable pins on this driver, desired speed control over the wheels is be obtained. This speed control is useful particularly during rotation of rover to properly align itself in the direction of the detected object.

Later, using an IMU sensor, the necessary calibrations for the required speed to overcome a terrain at a certain orientation are obtained. When test runs have been made on different surfaces, the orientation of rover in accordance with the surface and the speed with which the rover traveled is monitored. The accumulated error of the IMU has to be accounted for when analyzing the values.

An Optical encoder fixed at a wheel can be used if one desires to calculate the distance traveled by the robot. Necessary exception handling cases must be made so that the distance is not accounted while the rover is turning left or right. This data can be used to study the battery consumption for the distance traveled and help plan the length of the tasks accordingly.

B. Robotic Arm:

Necessary programming has been done for the Robotic arm to pick the detected object. For such manipulation of objects, Inverse Kinematics play a key role. The Robotic Arm movement occurs in two steps.

The rotation of the base takes first, followed by the rotation of the shoulder and elbow so that the end effector meets the target object.

When the object is detected, a bounding box is plotted and the center coordinates of that box are used as target coordinates.

First, the base (which helps is left-right motion) is set at 90 degrees. After the object is detected, considering from a top-down perspective, the inverse Tangent formula is used to obtain the required angle of rotation for the base. A small correction factor has been added after several test-runs using the "Bisection Method" to achieve better accuracy.

Once the Robotic Arm is set in the line of object, the gripper is opened and using inverse kinematics formula (obtained using the intersection of the circle formed by Links of the Arm), the angle of rotation for the elbow/actuator is obtained which helps it move in an up-down motion. As the rover stops at a fixed distance from the Target object, the angle of rotation for shoulder (which assists in front-back motion) is set default.

Thus, the gripper meets the target object and grabs it on and retrieves it back. For the smooth movement of the servos at the joints of the robotic arm, the PWM pulses sent to the servos are carried in steps instead of directly passing the final voltage value.

C. Object Detection:

For the demonstration, only two objects have been used. Nevertheless, with the use of MobileNet-v2, the network can classify images into 1000 object categories. Hundreds of images of both White and Black chess coins have been taken in different positions, environments, lighting conditions and these images were used to train the classifier using the SSD MobileNet-v2 model.

208 images were taken, of which 108 were of white chess coin and 100 were black chess pieces. Using labelImg, the images were labeled with bounding boxes, and these annotations are saved as XML files in the popular Pascal VOC format, which are converted into a single CSV file and then create the TFRecords (.record). Classes have been listed in the Label Map (.pbtxt). 168 were used as training data and 40 were as testing data. We used the ssd_mobilenet_v2_coco model for training as we were more concerned with the speed with reasonable accuracy. The training was done on the google cloud. For over an hour/22k steps with a batch size of 24, the training was run, but good results were obtained within 45minutes. Even though we have used only 2 objects for training, many more objects can be trained. The more data used for training; the better accuracy is obtained.

After picking the object, human face Detection is achieved using a pre-trained network in OpenCV called Haar Cascade Classifier. Among the several types of trained models available in the Haar cascade, the human face is used as detection criteria. Hence, the corresponding face detection file- 'haarcascade_frontalface_default.xml' is loaded before starting the detection. Then, the input of the camera is passed into 'detectMultiScale' functionality for the presence of a face.

VII. Working Mechanism

The program starts by initiating the webserver on raspberry pi. Different port Id's have been assigned to control the various flow of information. Initially, a web server with port id 8080 is activated which sets the rover in manual control mode. In this mode, the user can control the model by moving the rover manually using the joystick present in the app or controlling the Robotic arm's movement. The distance of the object against it can also be obtained on the App monitor. This mode also allows the user to capture photos and videos on command while hovering the rover.

To make the rover switch from manual to autonomous mode, port id 8081 is activated. The autonomous sequence is started when the user gives the voice command. The voice command should specify the object to be searched and collected by the rover. Now that the voice command is given, the system checks if the given command matches any of the objects it has been trained.

Currently, the system is trained for the detection of the Chess Pieces- White and Black Kings. Once the given command matches the object that the system is trained with, the rover starts hovering around to detect the object. The trained convolutional neural network is used for the detection of the specified object. Once the object is detected, using the center of the bounding box, the rover aligns itself in the line of motion to reach the detected object. The rover stops at 11.3cm away from the object. Robotic Arm collects the object using the coordinates of the object.

After picking the object, the Robot then starts searching for the presence of a human face using the Haar Cascade Classifier by pivoting at its place. Once the face is detected, the robot aligns itself and goes towards the human and drops the object at him.

In case of any emergency, the rover can be immediately be controlled manually, thus bringing in great flexibility of use.

VIII. Road Blocks during Testing

There were few setbacks when trying to make the model work.

- The toothed gears of the DC Motor lost their quality over-time due to usage. The primary reason for it was due to the overload on the rover, which had their impact on the wheels carrying the weight and hence the DC Motor which actuate the wheels. Also, most hindrances in the movement occurred when turning the Rover i.e. towards either left or right.

- There was a time delay when using the Ultra-sonic sensor. It was difficult to synchronize the functioning of the rover in accordance with the Calculated Distance. As a result, the rover would not start or stop at the desired distance. After many experiments, the Sync was established.

- The Robotic Arm could not be positioned properly on the rover. As a result, when the Robotic Arm would stretch far wide, it would end up hitting the camera module.

- Different algorithms were tested for Object Detection to find a balance between, less heat dissipated due to overuse of computational power and achieve higher frames per second. Neglecting this issue would lead to the breakdown of the working process on raspberry pi, due to overheating.

IX. Limitations

Despite trying the best to rectify the errors, there are a few limitations under which the device has to be operated.

- The object to be collected has to be placed at a height of above 22.3cm. Using the Rocker Bogie Mechanism demanded the Arm to be placed at a fixed height. So, the Robotic Arm cannot reach any object placed below 22.3cm, from the ground.

- The battery is the heaviest component used in the design, so when the rover is used for climbing any inclined surfaces, the battery pulls back the rover. So, it has to be ensured that the terrain is not too steep.

- If the object picked by the Robotic Arm is long along its length then the vision of the camera would be obstructed after that object is picked.

- The joints of the Robotic Arm, limit the action of Arm's movement. So, the joint has to be neither too tightly fitted nor too loosely held.

- For the delivery of the picked object, a human face must be in the field of vision of rover. Else the rover keeps rotating at its place until it finds one.

X. CONCLUSION

In Conclusion, a prototype of a multi-purpose rover has been designed. The scope of application for this project is versatile, not limiting its functionality to a fixed application. It can be used as an exoplanetary rover, to explore unknown environments and analyze it in real-time. This project can help the elderly, to get household objects on command. It can also aid the agricultural sector by analyzing the crop field and categorizing the ripe fruits or vegetables from that of spoiled ones and placing them in separate baskets. The limits of this project's application are one's thinking capacity.

XI. FUTURE SCOPE

- Eliminate the use of an Ultra-Sonic sensor and use a Camera for calculating the distance.
- Development of a reliable path navigation technique that can be employed on different surfaces.
- This is only a prototype that can be employed across various fields. Next, we would plan to develop a large-scale version of this project designed to help laborers in the construction industry to lay bricks for building wall autonomously without hindering the laborer's work. Also, the employment of this model in agricultural lands to retrieve ripe fruits without the farmer's assistance will be planned out.

XII. References

[1]. Alka N.U., Salihu A.A., Haruna Y.S. and Dalyop I.A. (2017), "A Voice Controlled Pick and Place Robotic Arm Vehicle using Android Application", American Journal of Engineering Research(AJER), Volume-6, Issue-7.

[2]. Bishwajit Pal, Samitha Khaiyum (2019), "Autonomous Planetary Vehicle Development Platform", International Journal of Engineering and Advanced Technology, Volume 9 Issue 1.

[3]. Chenguang Yang, Huaiwei Wu, Zhijun Li, Wei He, Ning Wang, Member, and Chun-Yi Su (2017), "Mind Control of a Robotic Arm with Visual Fusion Technology", Transactions on Industrial Informatics, Volume-3, Issue-4.

[4]. Daniel Anojan Atputharuban, Mobarakol Islam, Ravikiran Ramesh, and Hongliang Ren(2019), "Real-Time Instrument Segmentation in Robotic Surgery Using Auxiliary Supervised Deep Adversarial Learning", IEEE Robotics And Automation Letters, Volume 4, Issue 2.

[5]. Edward Tunstel (2017), "Operational Performance Metrics for Mars Exploration Rovers", Journal of Field Robotics, Volume 24, Issue 8.

[6]. Garima Aggarwal, Kevin Yu, and Debdeep Banerjee(2017), "Object Tracking Test Automation Using a Robotic Arm ", IEEE Access Journal, Volume6, Issue 5.

[7]. Harmeet Singh, Sangeeta, and Harpreet Kaur, "Design and Analysis of Robotic Rover with Gripper Arm using Embedded C", International Journal of Computer Applications Technology and Research, Volume 4, Issue 2.

[8]. Kyohei Otsu, Masahiro Ono, Thomas J. Fuchs, Ian Baldwin, and Takashi Kubota(2016), "Autonomous Terrain Classification With Co- and Self-Training Approach, IEEE Robotics And Automation Letters, Volume 1, Issue 2.

[9]. Mark Maimone, Ashitey Trebi-Ollennu, Edward Tunstel, Jeng Yen, Rich Petras, Reg Willson, "Mars Exploration Rover Mobility and Robotic Arm Operational Performance", IEEE Access Journal, Volume 2, Issue 10.

[10]. Piotr Czaplicki, Maciej Recko, Justyna Tołstoj-Sienkiewicz, "Robotic Arm Control System for Mars Rover Analogue", IEEE Robotics And Automation Letters, Volume 1, Number 16.

[11]. Tian Wang, *Member,* Yuting Yao, Yang Chen, Mengyi Zhang, Fei Tao, Hichem Snoussi(2018), "Auto-sorting System Towards Smart Factory based on Deep Learning for Image Segmentation", IEEE Sensors Journal, Volume 7, Issue 6.

[12]. Yasir Ali Memon, Imaaduddin motan, Muhammad Ali Akbar, Sarmad Hameed, Moez ul Hasan(2016), "Speech Recognition System for a Voice Controlled Robot with real-time Obstacle Detection and Avoidance", International Journal Of Electrical, Electronics And Data Communication, Volume 4, Issue 9.

[13]. Rasika Phadnis, Jaya Mishra and Shruti Bendale (2018), "Objects Talk - Object detection and Pattern Tracking using TensorFlow", Second International Conference on Inventive Communication and Computational Technologies.