Voice controlled wireless robot

Priyanshu Barua (CSE, 111815029)
Shubham (CSE, 111815037)
Shreyansh Gupta (CSE, 111816034)
Raushan Kumar (CSE, 111816029)

Guided by: Dr. Tanmoy Hazra

Outline

Our project is based on machine learning and robotics. It can be broadly divided into the following subprojects

1. Building a machine learning model for controlling the robot through speech.
2. Setting up the movement controls of the robot in the raspberry pi.
3. Connecting the raspberry pi and the machine learning model through socket programming for wireless control of the robot.
4. Integrating all the above parts to build a fully functioning robot.

Introduction

A robot is a preprogrammed machine capable of carrying out various tasks. Robots can be autonomous or semi autonomous.

Machine learning is a series of algorithms and concepts that allows computers to learn and perform tasks that require human intelligence.

In this project we have built a basic semi autonomous robot which uses machine learning to identify commands given to it.

Motivation

The motivation behind our project is to learn the machine involved in voice controlled robots. This will help us to build more advanced robots that can be used wirelessly for searching areas hit by catastrophes. Thus, help save people stuck in the rubble and unnoticed places during natural disasters or accidents like collapsing of buildings.

They can also be used as delivery robots in offices, hotels and schools.

Literature Survey and Research Gaps

1. State of the Art Speech Recognition – Google, USA
2. Experimental Speech Recognition using Raspberry Pi – Hasan Gyulyustan, Svetoslav Enkov

State of the Art Speech Recognition using Sequence to Sequence Models – Google, USA

Literature Survey

1. This paper talks about a speech recognition model based on the Listen - Attend - Spell architecture.
2. The listener converts the speech into higher level feature representation, the attender learns the alignment between the input features and the speller converts the attender’s output into text.
3. The three models are trained together.

State of the Art Speech Recognition using Sequence to Sequence Models – Google, USA

Research Gaps

1. This model cannot process speech in real time.
2. This model has only been trained in 22 million audio - text utterances. Usually, models are trained in much larger data.
3. This model is not able to use proper spellings for relatively lesser used words.

Experimental Speech Recognition based on Raspberry Pi – Hasan Gyulyustan, Svetoslav Enkov

Literature Survey

1. This paper first gives a brief overview of the various speech recognition models currently in existence like Dynamic Time Warping model, Hidden Markov model, Listen-Attend-Spell model.
2. Implements Google Speech Recognition API in Raspberry Pi.

Research Gaps

1. The voice input has to be direct. Cannot take voice input from a phone.

Problem Statement

To build a robot that takes commands through speech wirelessly and performs accordingly. This will involve the following.

- Building a machine learning model for command recognition.
- Forming a wireless connection between the robot and the controlling device.
- Integrating and fine tuning the robot so that it’s functioning is practical.

Objective

The objective behind this project is to do the following

1. Learn machine learning and build a speech recognition model with the help of it.
2. Implement the above model in a Raspberry Pi so that the robot can function remotely.
3. Learn how to wirelessly connect the robot and the voice sender.

Methodology

1. Command Recognition Model.
2. Wireless connection between the robot and voice sender.
3. Making the robot

Command Recognition Model.

1. For the robot to identify the commands and function accordingly, we have used a machine learning model.
2. It is a simple classification model built on a self made data set.
3. The Google recognizer converts the input speech into text which in turn becomes the input for this model.
4. The model converts the text into processable data and uses a DNN to identify the command.

Wireless connection between robot and voice sender.

1. To wirelessly connect the voice sender(laptop) and the robot, we have used socket programming.
2. Socket programming facilitates information exchange between a server and a client. In this case, the laptop acts as the server and the robot as the client.
3. The command recognition model produces an integer that signifies a particular command. This integer is sent to the robot through the socket.

Making the robot

1. We have used a raspberry pi as the processor for the robot.
2. The integer sent through the socket is received by the raspberry pi which implements the code for the robot's movement.
3. For it's movement the raspberry pi controls the motor driver which in turn controls the two motors for the wheels.
4. We have used a power bank for powering the robot.

Results and Discussions

At the end of this project we were able to achieve the following

1. A working command recognition model with an accuracy of 95% in 600 epochs.

```
[A[A Training Step: 3600  | total loss: [1m[32m0.27675[0m[0m | time: 0.083s
[2K| Adam | epoch: 600 | loss: 0.27675 - acc: 0.9508 -- iter: 45/45
```

2. A semi autonomous robot that can be wirelessly controlled through speech with real time video projection.

Observations and Challenges

In the due course of the project we faced several challenges but because of our interest and hard work we were able to overcome most of them. Some of our observations are

Machine learning is a very fascinating part of computer science with lots of development potential.

The raspberry pi is a new generation mini computer with a lot of amazing applications.

Conclusions

We obtained the following conclusions

We were able to achieve our initial objective of introducing ourselves to machine learning and implement the learnt concepts in a robot.

We learnt the basics of raspberry pi and socket programming.

Last but not least, with close deadlines and lots to do we learnt the importance of time management.

Future Scope

1. Using Picamera for face recognition.
2. An autopilot offline mode, in case the connection between the robot and controller is lost.
3. Using sentiment analysis to identify people in crisis.
4. A better and more efficient design.

References

For building the project we went through a number of research papers, some of which are listed below

1. State of the Art Speech Recognition – Google, USA.
2. Experimental Speech Recognition using Raspberry Pi – Hasan Gyulyustan, Svetoslav Enkov. Published in the IOSR Journal in 2017.
3. English Conversational Telephone and Speech Recognition by Humans and Machines – IBM Research, Tokyo, Japan.
4. Speech Commands: A dataset for limited vocabulary speech recognition – Pete Warden, Google Brain, California, USA.