Gesture Based Robot with Victim Detection using Audio

Anju Varghese[1], Vaishak Nair[2], Aathira Lakshmi Pillai[3], Sreelakshmi Pillai[4], and Prof. Suresh Babu[5]

[1,2,3,4,5]*Department of Computer Engineering, Pillai College of Engineering, Navi Mumbai, India – 410206*

Abstract - *When a large-scale disaster such as a huge earthquake occurs, the buildings collapse and many people get buried under the rubble. Under such circumstances, it is very important to detect such casualties effectively in order to promptly rescue the victims. It is difficult for a human rescue team to get under such small gaps to help the trapped victims. A compact robot with collision avoidance system, on the other hand, can easily navigate through the gaps. In our project, the proposed robot will be able to fit into confined spaces, detect obstacles, capture audio and detect human presence by filtering out background noises and further processing the signal to extract human voice. This extracted signal will then be further compared with a speech database and Keyword Spotting technique is employed to detect whether the uttered word belongs in the set of predefined keywords.*

Keywords—Voice and Gesture Controlled Robot, Human Detection, Object Detection, Voice Recognition, Audio Processing

1. INTRODUCTION

1.1 Fundamentals

Robot Methodology

A robot due to its compact size and mobility can access areas that are beyond the scope of human controllers. Traditional robots used wired controllers that drastically reduced its range of motion. Using wireless technology the range can be exponentially increased. The gesture control aspect of the robot is governed by a wifi module(RF443) that transmits the data produced by the accelerometers mounted on the gesture controller.

Band-Pass Filter

The band-pass filter acts as the first source of audio processing. In the proposed areas of operation for the robot, the background noises can muffle out the voices of the victims. To counter this a band-pass filter is deployed right after audio acquisition which filters out all audio beyond the human voice audio range.

Keyword Spotting Using DNN

The processed audio received is searched for keywords using keyword spotting. Deep neural networks are employed for the same. The tensorflow node of the input signal is added to the DNN that detects and predicts the keywords. This is done by training the network based on the desired keyword that should be detected. The output of the neural network is the part of the signal with only the desired keyword present. This is used to detect human presence in the environment.

1.2 Objectives

1. To successfully control the movement of the robot using gesture and voice commands.

2. To capture and extract audio using the microphone attached to the robot.

3. To process and filter the extracted audio signals to remove noise and other interferences using band pass filters.

4. To recognize the audio signal by comparing it against a speech database like Julius, using Keyword Spotting technique.

5. To capture the location of the victim, based on the GPRS sensor.

6. To aid in search and rescue activities during natural calamities.

2. LITERATURE SURVEY

A. Accelerometer and Arduino Based Gesture Controlled Robocar: The paper summarises the creation and deployment of a 4 wheeled bot based on gestures, voice commands and auto drive. The RF module is a cheap transmitter receiver pair the sole purpose of which is to transmit the data of the accelerometer to the arduino. The glove contains the accelerometer. The transmitter receiver is coded into the arduino to receive and transmit processed data. The self driving mode of the bot used 3 ultrasonic sensors and 1 servo motor. servo motor. The HC-SR04 module is responsible for processing the information from the sensors. Depending on the minimum distance that is set up in the Ardino the sensor on the motor would move in 90 degrees fashion and calculate a distance larger than the minimum distance if an object is detected with imminent collision.

B. Robot Controlled Using Hand Motion Recognition: The paper is concerned with the creation of a gesture controlled 4 wheeled robot that also has an arm attached to it for grabbing action. The movement aspect of the bot is controlled by the accelerometers which is responsible for noting the changes in the actuated values of the 3D accelerometer and transmitting the necessary signals. The zigbee is a low powered unit that is used to create personal wireless networks for transmission of data and signals. Due to its low power consumption the effective range of a zigbee is 100 metres. the ARM-7 is a microcontroller that converts the signals generated by the accelerometer into data signals and sends it to the zigbee for transmission. At the receiver side the ARM-7 takes the incoming data packet, process it and take the necessary actions. The flex sensor is used to control the robotic arm. Depending upon the degree of bending of the flex sensor the corresponding values are transmitted which is processed to control the movement and the clawing action of the arm The motor driver circuit works on a H-Bridge principle in which the current can be redirected to either branch of the circuit thereby permitting the spot turning of the robot by activating only parts of the circuit necessary for the turn.

C. Victim Detection Using UAV Victim Detection Using UAV with On-board Voice Recognition System: This paper proposes a novel victim detection system using a drone. The drone is controlled using Navio2 chipset.Navio2 is also capable of audio processing. The Speaker emanates a certain word/audio which is heard by the victims and they respond accordingly. The response of the victims received is mixed with a lot of unnecessary background information. Extraction of human audio from the input signal is done by implementing various Band Pass Filter (BPF) programmed in python. Julius is a software that is used for audio processing. Based on the response of the victim the response word is checked against a database of such words and the information along with the location of the victim is saved into a text file using GNSS.

D. Design and Implementation of Sound Searching Robots in Wireless Sensor Networks: In this paper, a sound target-searching robot system is implemented. It consists of a 4-channel microphone array for sound collection magneto-resistive sensor for declination measurement, and a wireless sensor networks (WSN) for exchanging information. The system features of Digital Signal Processing and improved spectral subtraction method is used for noise reduction Based on the neighbor classification method, we match the trained feature template to recognize sound signal type. Improved cross correlation given as input to microphone arrays for sound recognition by implementing sound localization and searching using magneto resistive sensors. FIR band-pass filter is used to improve the traditional GCC method.

3. PROPOSED WORK

This project aims at developing a system that is controlled by gesture commands and can successfully detect the presence of humans using audio signals. This chapter explains in detail, the existing architecture, its functioning and components as well as the proposed architecture and the techniques to be used for implementing the same.

3.1 System Architecture

The system architecture is given in Figure 1. Each block is described in this Section.

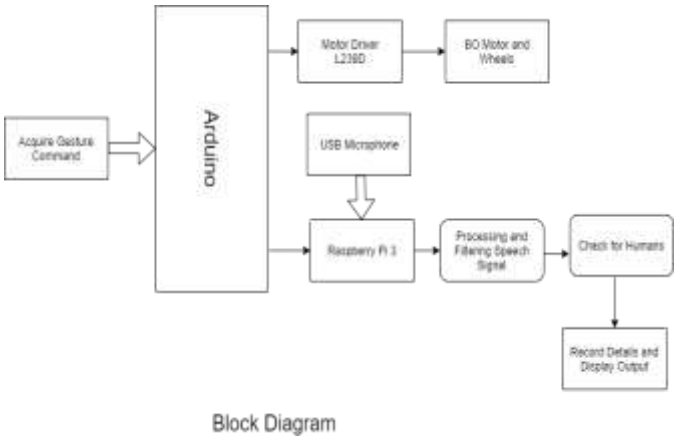

Fig. 1 Proposed system architecture

A. Acquire Gesture Command Block Description: Here, the input is given to the Arduino microcontroller by the ADXL335 Accelerometer which is attached onto a glove that is worn by the user to generate gestures. The accelerometer measures gravity on the 3-axis. This is given as input to the Arduino.

B. Based on this input, the Arduino is programmed to move the motors and wheels, by signalling the motor driver L239D. In this way. The movement of the robot is controlled using an accelerometer and Ardunio.

C. Audio Processing: The audio processing part consists of two major components i.e cleaning the incoming audio and keyword spotting. In order to clear out any other signals that may disrupt the processing part a band pass filter is applied. The band pass filter selects only those audio ranges that are generated by human sources and clears out the background noises. A band pass filter is preferred over other modes of noise processing due to its simplicity in implementation and low computational cost. Complex algorithms would drastically increase the dependency of the robot on high powered systems thereby decreasing its effective deployment range. The processed

audio is then sent to a keyword spotting deep neural network. This network would be trained to identify and predict keywords coming from the audio. A deep neural network is preferred over HMM due to its increased accuracy and faster implementation.

D. Location Detection: Upon receiving the desired keyword from the DNN, the network signals the onboard Navio2 chipset that activates the GNSS module and writes the location of the source of audio onto the onboard SD chipset or transmits it live to the rescuer as a text file.

4. REQUIREMENT ANALYSIS

The implementation detail is given in this section.

4.1 Software

The robot is capable of detecting voices present in confined space. The voice or signal that has been detected is acquired by the robot. This signal undergoes audio processing and filters out human voices and removes all the disturbances present in the actual signal. Now, from the human voice, certain keywords like "help", "I am stuck", etc. are spotted using DNN techniques.

4.2 Hardware

The robot is attached to several hardwares for providing various functionalities. Arduino UNO, a microcontroller is used for input and output operations. With so many pins available, you can easily read data from sensors, or control different motors and actuators. Accelerometer is used for measuring acceleration tilt angles by using the MEMS (Micro-Electro Mechanical Systems) technology, where acceleration is the rate of change of velocity with respect to the time. Motor drivers act as an interface between the motors and the control circuits. Encoder and decoder are used to convert the data from one form to another form.

ACKNOWLEDGMENT

It is our privilege to express our sincerest regards to our supervisor Prof. Suresh Babu for the valuable inputs, able guidance, encouragement, whole-hearted cooperation and constructive criticism throughout the duration of this work. We deeply express our sincere thanks to our Head of the Department, Dr. Sharvari Govilkar and our Principal, Dr. Sandeep M. Joshi for encouraging and allowing us to present this work.

REFERENCES

1. Lavanya K N, Ramya Shree D, Nischitha B R, T Asha, C Gururaj, "Gesture Controlled Robot", 2017 International Conference on Electrical, Electronics, Communications, Computer and Optimization Techniques (ICEECCOT) pp.2017 Available:[Online] ieeexplore.ieee.org/abstract/document/8284549.

2. Humayun Rashid, Iftekhar Uddin Ahmed, Sayed Bin Osnan, Qader Newaz, Md. Rasheduzzaman and S M Taslim Reza ,"Design and Implementation of a Voice Controlled Robot with Human Interaction Ability", 2017 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering pp.2017.

3. Yuki Yamazaki, Masaya Tamaki, C. Premachandra, C. J. Perera, S. Sumathipala,

   B.H Sudantha, "Victim Detection using UAV with On-board Voice Recognition System", 2019 Third IEEE International Conference on Robotic Computing pp.2019.

4. Arkhaprabha Lodh, Debopama Ghosh, Debosama Ghosh, "Accelerometer and Arduino based Gesture Controlled Robocar", 2018 International Journal of Innovative Research in Science, Engineering and Technology pp.2018 Available:[Online] http://www.ijirset.com/

5. Premangshu Chanda, Pallab Kanti Mukherjee, Subrata Modak, Asoke Nath, "Gesture Controlled Robot using Arduino and Android", 2016 International Journal of Advanced Research in Computer Science and Software Engineering Research Paper pp.2016.

   Available:[Online] www.ijarcsse.com

6. Gabriel Oltean, Lăcrimioara Grama, Laura Ivanciu, Corneliu Rusu, "Alarming Events Detection Based on Audio Signals Recognition", 2015 International Conference on Ubiquitous Robots and Ambient Intelligence (URAl) pp.2015.