A Project Report

on

Speech Controlled Robot

(Technical Report)

Submitted by:

Akhilesh Kumar – 174101020

Pammi Sairam – 174101014

Md Zaki Anwer – 174101030

Moirangthem Kirshnananda Singh - 174101005

Under the esteemed guidance of

Prof. P.K.Das

SPEECH CONTROLLED ROBOT

ABSTRACT:

This report is about our project speech control robot. This report contain the brief overview of our application like its need, implementation details and the results of this project.

INTRODUCTION:

Speech Controlled Robot is a robot which can be controlled by human voice. Human voice will control the motion of robot by issuing some voice commands like: move, back, left, right and stop. There is a 2d graphical view of this robot on computer which will move on basis of given voice command.

People want to just give voice command instead of doing something physically, such as reading mails, using house appliances like microwave ovens, washing machine etc.
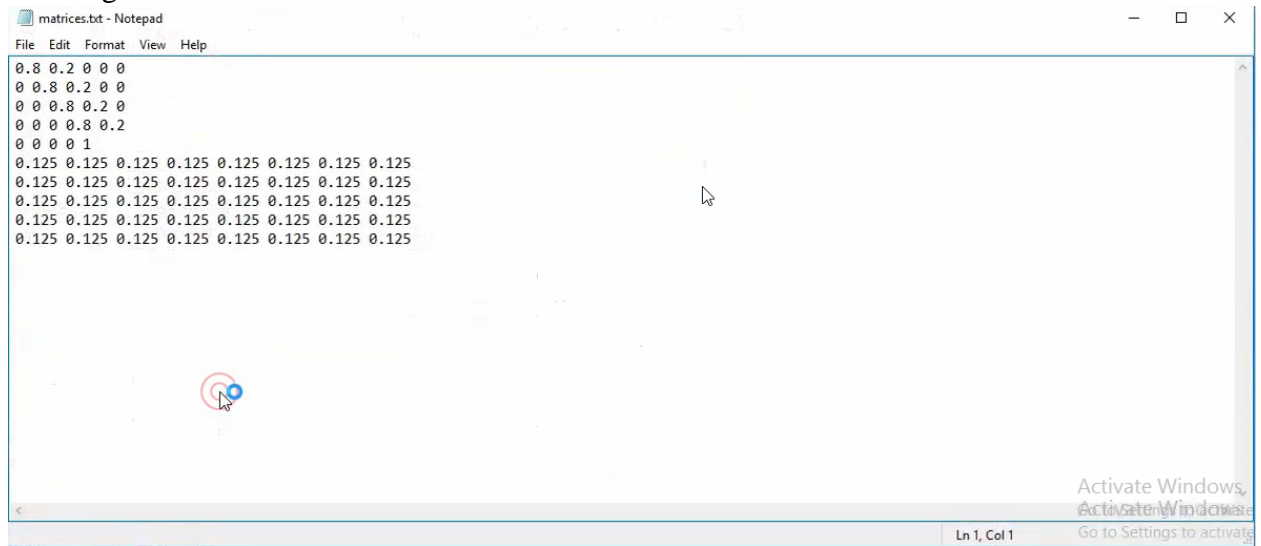
IMPLEMENTATION DETAILS:

1. Generation of codebook: we are training codebooks from the universe of cepstral coefficients of vowels using LBG. First we calculated the centroid of universe and splitting it into two vectors by adding and subtracting a very small value. Then converging these vectors and again splitting until we get the desire number of cells in codebook.
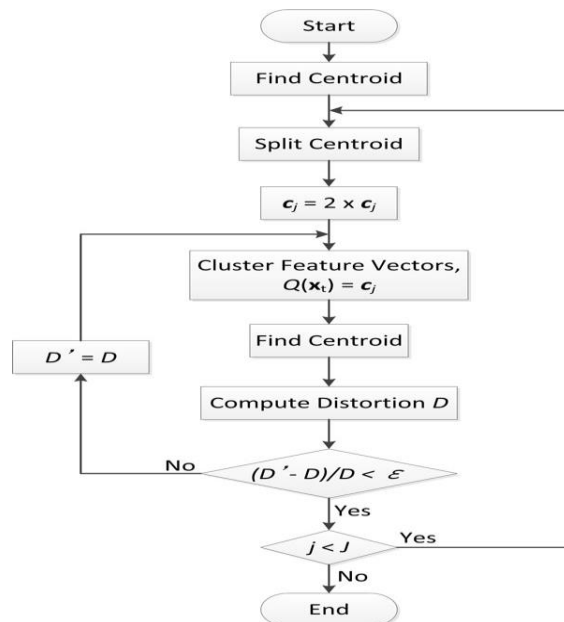
2. Collection of training data: We have collected 20 utterances each of move, left, right, back, stop and trained HMM model corresponding to each word. Here we used our previously generated codebook to train model. We used 5 states Bakis model and updating the parameters in each iteration of training to get more accurate model.

3. We have made gui of this application in visual studio using graphics library of winbgi WinBGIm is a graphics library of C++ to implement graphics in visual studio using C++.

4. For testing we saved the voice in computer using microphone and tried to find the observation sequence using the previously generated codebook. Then applied forward procedure to find the probability that this observation sequence belongs to the given model for each distinct model which we have already generated. The model which gives the highest probability is the recognized word on which this model is trained. Then we issued the command for motion of robot.

SNAPSHOTS:

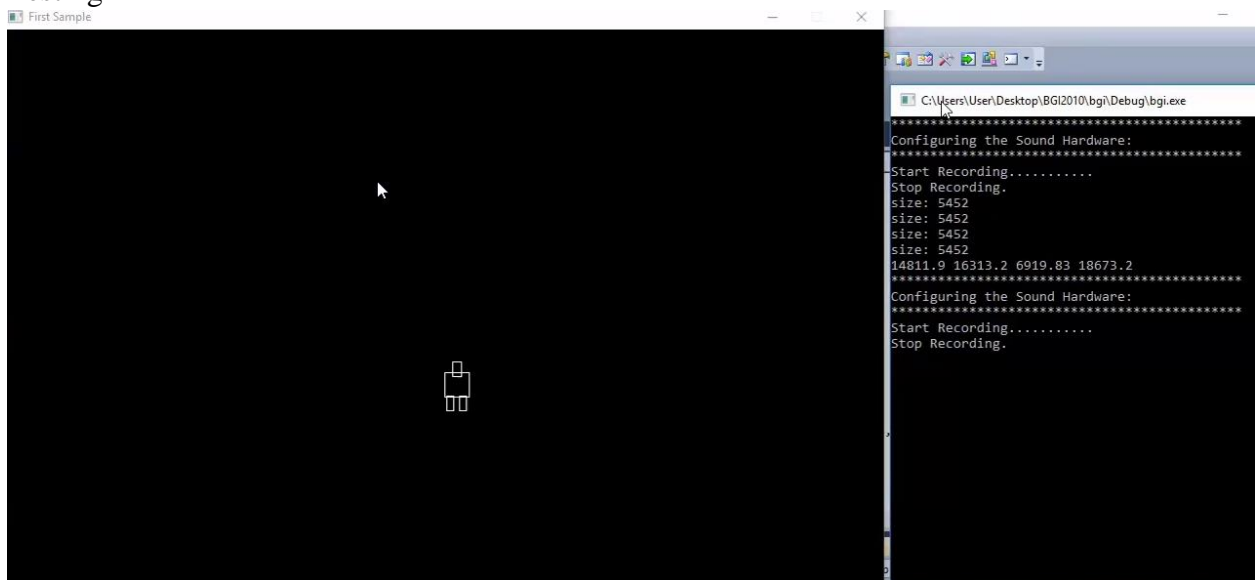1. Training

```
0 0 0 0 1.0001645390544753

B matrix:
9.9999999999999996e-081 9.9999999999999996e-081 0.00026047107055946655 9.9999999999999996e-081 1.1120129103223976e-006 4
.8654380828563996e-007 0.060333768513272747 0.9394041618594493
9.9999999999999996e-081 9.9999999999999996e-081 0.0038976426627848875 9.9999999999999996e-081 3.4449340799084084e-005 1.
6773147922078439e-005 0.25863955917080428 0.73741157567768911
9.9999999999999996e-081 9.9999999999999996e-081 0.019222033199093682 9.9999999999999996e-081 0.00037486161553431285 0.00
015092960389843625 0.29949010747339683 0.68076206810807649
9.9999999999999996e-081 9.9999999999999996e-081 0.059676213878844565 9.9999999999999996e-081 0.0020224114087927104 0.000
739925171396071825 0.15778594083117303 0.77977615216722962
9.9999999999999996e-081 9.9999999999999996e-081 0.034397480749967372 9.9999999999999996e-081 0.04517607978893888 0.037691
833298799347 0.1895765433816983 0.69315806278088954

A matrix:
0.7727210043658268 0.20126759288878104 0 0 0
0 0.78747495013387592 0.28089857577458344 0 0
0 0 0.7981092712272243 0.18977104738607897 0
0 0 0 0.89461682238622886 0.15801851972533049
0 0 0 0 0.99489215485687166

B matrix:
9.9999999999999996e-081 9.9999999999999996e-081 1.5060713015466918e-007 9.9999999999999996e-081 2.4010037868357612e-025
3.422101054305581e-047 0.0093094476285112224 0.99069040176435841
9.9999999999999996e-081 9.9999999999999996e-081 0.00012425569585816431 9.9999999999999996e-081 3.6816000799661753e-013 7
.049408259226797e-026 0.3246738377057464632 0.67520190659802826
9.9999999999999996e-081 9.9999999999999996e-081 0.0032655502814787794 9.9999999999999996e-081 8.5140790134967514e-008 3.
037207360338983e-016 0.38480870293714103 0.61192566164058937
9.9999999999999996e-081 9.9999999999999996e-081 0.16478260953289234 9.9999999999999996e-081 3.8615369015257142e-005 3.85
```
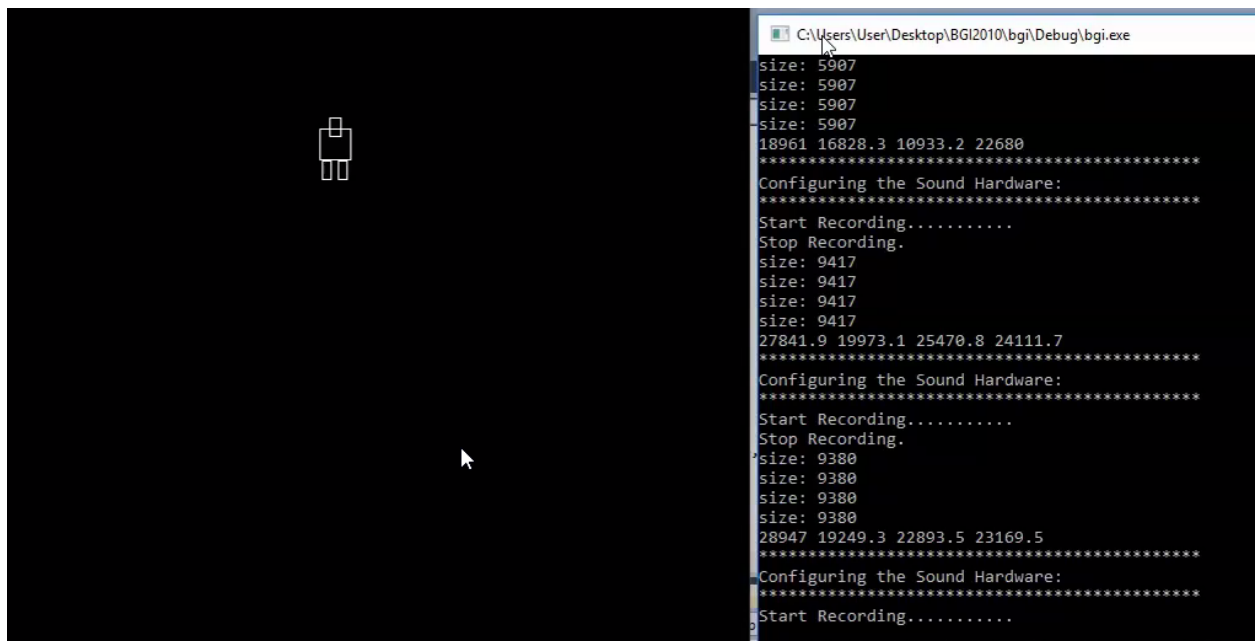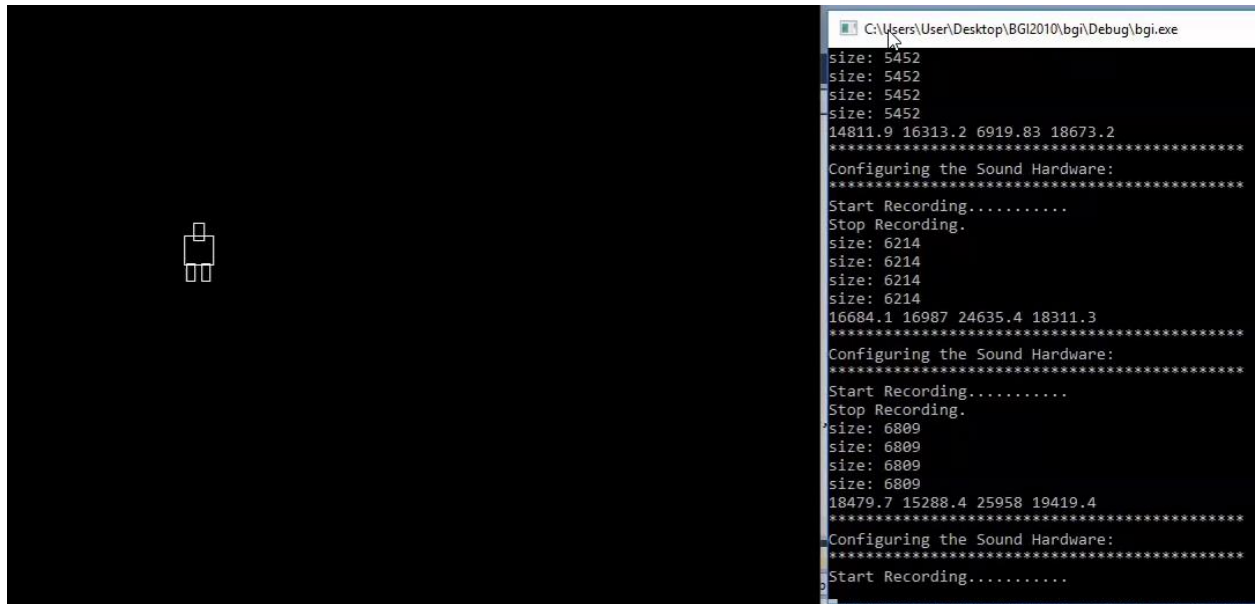
File Edit Format View Help

```
0.7884749046182804 0.21542502241574751 0 0 0
0 0.77401954675855456 0.22209643181087416 0 0
0 0 0.80241424237650083 0.22050752200120455 0
0 0 0 0.85313414200773785 0.14167550581891286
0 0 0 0 0.99955922097735372
9.9999999999999977e-081 9.9999999999999977e-081 0.062824003416818486 0.0013174544535758921 0.3073056749708305 0.027697680974872622 0.00093094476;
9.9999999999999977e-081 9.9999999999999977e-081 0.063426349970870183 0.019951726033722826 0.20520634968404497 0.11666531283333961 0.032467413465;
9.9999999999999977e-081 9.9999999999999977e-081 0.13253668819784103 0.0016397314556954724 0.095111217676097787 0.069091407973862234 0.0386268603;
9.9999999999999977e-081 9.9999999999999977e-081 0.20754025150339536 1.4706619997050809e-006 0.0095695651192202226 0.02502214664449337 0.02912106;
9.9999999999999977e-081 9.9999999999999977e-081 0.11320325757541314 0.0041735531560887001 0.0065871241487334073 0.017653559921573855 0.163005567;
```

2. Testing

```
**********************************
Configuring the Sound Hardware:
**********************************
Start Recording..........
Stop Recording.
size: 5452
size: 5452
size: 5452
size: 5452
14811.9 16313.2 6919.83 18673.2
**********************************
Configuring the Sound Hardware:
**********************************
Start Recording..........
Stop Recording.
```

Speech Controlled Robot

REFERENCES:

1. CS 566 Speech Processing Lectures, by Prof. P. K. Das, IIT Guwahati
2. https://en.wikipedia.org/wiki/Hidden_Markov_model

Speech Controlled Robot