

Implementation of a Speaking Calculator

Platform of development: C(language)
(Project Title)

Project report submitted in partial fulfilment of the requirement for the degree of

Masters of Technology

By Group members :-

1. **Abinash Deka(224101004)**
2. **Sanjoy Das (226156004)**
3. **Shikha (224101066)**

Under the Guidance of:

Prof. P. K. Das

Department of CSE, IITG



Department of computer science and engineering

INDIA INSTITUTE OF TECHNOLOGY GUWAHATI

(July-Nov 2022)

Abstract

The voice based calculator, that is used to perform basic arithmetic operations. It recognises the isolated spoken digits from 0-9 and basic arithmetic operations like add, subtract, multiply and divide. It then performs the respective arithmetic operations, and displays the final output on a screen.

CONTENTS

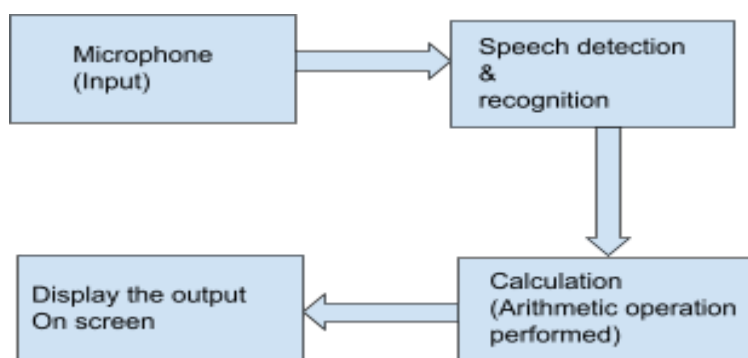
1. Introduction.....	4
2. Objective.....	5
3. Project features.....	6
4. Conclusion.....	9
5. Future work.....	10

Introduction

Speech recognition based technology is one of the fast growing engineering technologies. It has a number of applications in different area and provides potential benefits. In the world many people are suffering from various disabilities, and many of them are blind or unable to use their hands efficiently. For such cases the speech recognition system plays a significant role to help them.

Our approach to these issue is to provide real time visual feedback of data spoken. This project is designed and developed keeping that factor in mind. Our project “Speaking Calculator” is capable to recognise the speech and ability to perform simple arithmetic operations like add, subtract , multiply and divide.

At the initial effort is made to provide help for basic arithmetic operations as discussed above, but it can further be updated and enhanced in order to cover more arithmetic operations.



Block diagram of Speaking Calculator

Objective

The objective of the project is to design and develop a Speaking Calculator with single digit operations. The operations that can be performed are - Addition, Subtraction, Multiplication and Division. The project also includes training for new users as well as testing of the model and user. The project has been done towards partial completion of the course CS 566 -Speech Processing in July-Nov 2022 session.

Project Features

The various features available in the project are:

- Creation of user
- Training an user
- Testing an user

Creation of user

The project provides the facility to register new user and record user utterances for digits 0-9 and operations addition(+), subtraction(-), multiplication(*) and division(/) using the Recording Module. The recorded files are stored under the respective user directory under the application folder for easy reference to the files respective to an user.

Training User

We have used 30 speech utterances of the digits and operations for the training purpose. The training phase consist of the following operations for each utterance:

1. Read and stores the speech samples in an array
2. Calculation of mean noise
3. Next, we have calculated DC Shift for the speech samples by subtracting the mean noise from the samples.
4. Next, we have normalised the speech data to +/- 5000 (max), by initially finding the highest and lowest peak of the speech samples.
5. Next, for each of the 320 frames of samples (shifted by 80), we have calculate the:
 - Hamming Window
 - Computed Ri's

- Computed Ai's
 - Computed Ci's
 - Applied raised Sine window on Ci's
6. Exported the Ci's for each digit and operations and generated a universe of the same.
 7. Next, we have generated the codebook of size 32 using the universe and LBG algorithm.
 8. Next, we have generated the observation sequence for all the training dataset.
 9. After that, Modelling of data by using HMM(Hidden Markov Modelling) following:
 - Solution to problem 1
 - Solution to problem 2
 - Solution to problem 3
 by taking the initial a.txt, b.txt and pi.txt as input.
 10. After that we have generated averaged model for each of the digits and the operations which are to be used for the testing phase.

Testing User

In the testing phase, user inputs are recorded using the Recording Module as:

- Input utterance for the first operand (0-9)
- Input utterance for the second operand(0-9)
- Input utterance for the operation (+, -, *, /)

Based on the user input utterances, we have calculated/found out the observation sequence for both the operands and the operation, by following the above steps.

Using these three observation sequences, we have calculated solution to problem 1 of the HMM model by taking reference of all the averaged models of the training phase to predict the actual digit and operation.

Based on the recognized digit and operation, we have performed the specific mathematical operation and displayed the results for the user.

Conclusion

The system performed most of its functions at the end of the design. The speaking calculator was able to recognize spoken words in natural environments and was able to run the recognition code on it. The resulting output would get displayed on the screen in real-time. The time taken for the whole process was very acceptable. Due to the restricted size, the training codebook could not be made larger. Only 30 utteration per classification word were used in this system. But, a larger, more controlled training set could result in higher recognition accuracy.

We had a great experience while doing the project using the knowledge gathered during the course and we believe that others may be benefited through this project by taking it as a base for development of real time applications.

Future work

This project currently performs the recognition for single digit operand only. In future, we can implement more operations such that log, ln, square root, sin, cos and many more operations by modifying the code. We can also implement analogue voice output instead of the screen display so it can be used for blind people too.