ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Issue 3, March 2017

Speech SMS

Jayaram.B[1], Gokilapriya.M[2], Kiruthika.M[3], Mohana Priya.P[4], Thamaraiselvi.R[5]

Assistant professor, Department of computer science and Engineering, INFO institute of engineering, Coimbatore, India[1]
UG scholar, Department of computer science and Engineering, INFO institute of engineering, Coimbatore, India[2]
UG scholar, Department of computer science and Engineering, INFO institute of engineering, Coimbatore, India[3]
UG scholar, Department of computer science and Engineering, INFO institute of engineering, Coimbatore, India[4]
UG scholar, Department of computer science and Engineering, INFO institute of engineering, Coimbatore, India[5]

Abstract: Speech recognition was difficult to use, but with elevation in technology. Now, it is possible to generate the desired output. Speech SMS is an application, in which we can convert the voice input into the text using the speech recognition algorithm in offline. In the existing system, the phone number can be retrieved from the database, but in this project, the contact can be saved to the storage, through the voice input and it can only be retrieved from the database. The body of the message are given as speech input which is converted as text message by detecting the pitch rate, using the pitch detection algorithm. If the send button is clicked, then the message will be delivered to the receiver and it will be stored to the database. It reduces the time used in typing the content of the message by getting the speech input of the message.

Keywords: Speech recognition algorithm, Hidden Markov model, Pitch detection algorithm, Speech to text conversion

I. Introduction

The speech input can be used in varying domains such as automatic reader and for inputting data to the system. Speech recognition can minimize the use of text and other types of input, at the same time minimizing the calculation needed for the process. Decade back Voice or signalled input is inserted through any speech device such as microphone, then speech can be processed and convert it to text hence able to send SMS, also Phone number can be entering either by voice or may select it from contact list.

II. Related Works

Speech Recognition stands majorly on five pillars that are, feature extraction, acoustic models database which is built based on the training data, dictionary, language model and the speech recognition algorithm. The inputs data i.e. voice are first converted to digital signal and are sampled on time and amplitude axis. This digitalized signal is then processed. For processing the signal is divided into small intervals, which depends on the algorithm used. The generalized timestamp is 20 milliseconds. This division is based on the features of data as those features are compared with database element. Database element contains information of feature of the word found and according the command is created. The basic element can be a phoneme for continuous speech or word for isolated words recognition.

III. Scope Of Project

This project can be deployed as a mobile application or as an add-on system to the existing message sending system. “Speech SMS” can help us to send the message for which the input is given as speech. It minimizes time, instead of typing the messages. The phone numbers can be retrieved from the database using the speech input of the contact name. It can be done in offline for the processing of the speech recognition of the input speech.

IV. Proposed Methodology And Discussions

The system shown here will use SR with Google server which uses HMM method. The brief description of how speech is recognized is as follows. Firstly the speech is given as input speech can be fluctuating set of signals which are recorded. These signals depends on speaker how is his/her voice quality and hold on the language. The input data is divided into words and phrases, i.e. command is divided into several parts. Lastly comes the processing phase where accordingly system understands command and

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Issue 3, March 2017

execute.

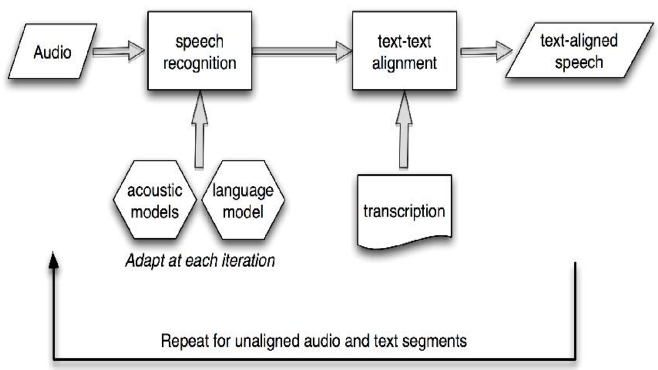

Fig.1. System Architecture diagram

The system here is divided into small subsystems. In our application the models are as follows. Firstly user has an option to select whether to use voice as an input or select contacts manually. If user selects manual option then a service is called which access all the contacts in contact list present on the cell phone. If voice is selected then the Google SR API is called and a dialog box appears which says that speck now and a MIC type image is formed. Once the user is done speaking then API takes a few seconds to process the data and output is displayed on the sender's address block.

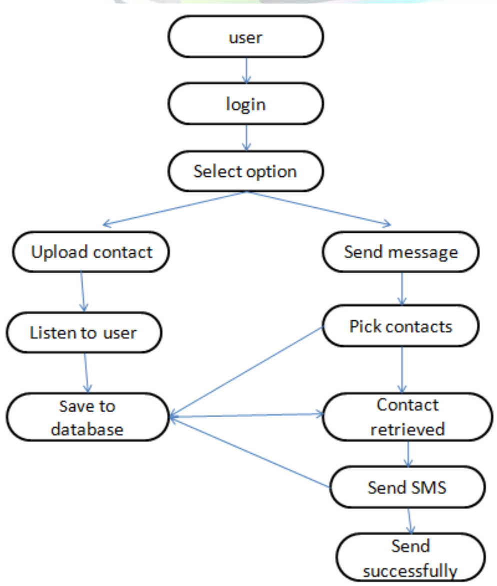

Fig.2.Flow diagram

A. Speech recognition:

Speech Recognition (SR) is the inter-disciplinary sub-field of computational linguistics that develops methodologies and technologies that enables the recognition and translation of spoken language into text by computers. It is also known as "Automatic Speech Recognition" (ASR), "computer speech recognition", or just "Speech to Text" (STT). It incorporates knowledge and research in the linguistics, computer science, and electrical engineering fields.

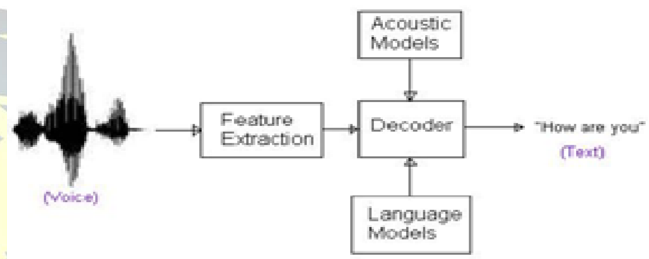

Fig.3.Speech Recognition system

Some SR systems use "training" where an individual speaker reads text or isolated vocabulary into the system. The term speech recognition or speech identification refers to identifies what the speaker says, rather than the voice. Recognizing the speech can simplify the task of typing the input message in systems.

Both acoustic modelling and language modelling are important parts of modern statistically-based speech recognition algorithms. Hidden Markov Models (HMMs) are widely used in many systems. Language modelling is also used in many other natural language processing applications such as document classification or statistical machine translation. The pitch rate of the speech input is detected using the pitch detection algorithm.

B. Hidden Markov models (HMM)

Modern general-purpose speech recognition systems are based on Hidden Markov Models. These are statistical models that output a sequence of symbols or quantities. HMMs are used in speech recognition because a speech signal can be viewed as a piecewise stationary signal or a short-time stationary signal. In a short time-scale (e.g., 10 milliseconds), speech can be approximated as a stationary process. Speech can be thought of as a Markov model for many stochastic purposes.

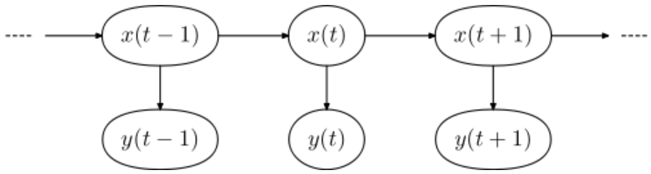

Fig.4.Hidden Markov Model

Another reason why HMMs are popular is because they can be trained automatically, simple and computationally feasible to use. In speech recognition, the hidden Markov model would output a sequence of n-dimensional real-valued vectors outputting one of these every 10 milliseconds. The vectors would consist of cepstral coefficients, which are obtained by taking a Fourier transform of a short time window of speech and decor relating the spectrum using a cosine transform, then taking the first (most significant) coefficients. The hidden Markov model will tend to have in each state a statistical distribution that is a mixture of diagonal covariance Gaussians, which will give likelihood for each observed vector. Each word, or (for more general speech recognition systems), each phoneme, will have a different output distribution; a hidden Markov model for a sequence of words or phonemes is made by concatenating the individual trained hidden Markov models for the separate words and phonemes.

C. Pitch Detection Algorithm

A Pitch Detection Algorithm (PDA) is an algorithm designed to estimate the pitch or fundamental frequency of a quasi-periodic or oscillating signal, usually a digital recording of speech or a musical note or tone. This can be done in the time domain or the frequency domain or both the two domains.

Pitch detection is very important for many speeches processing algorithm. Speech recognition system of tonal language use pitches tracking for speech recognition, which is important in disambiguating the myriad of homophones. Pitch is also crucial for prosodic variations in text-to-speech systems and spoken language systems. The principles of the two pitch detection algorithms, pre-processing and the extraction of pitch pattern techniques are introduced. One simple approach would be to measure the distance between zero crossing points of the signal (i.e. the zero-crossing rate).

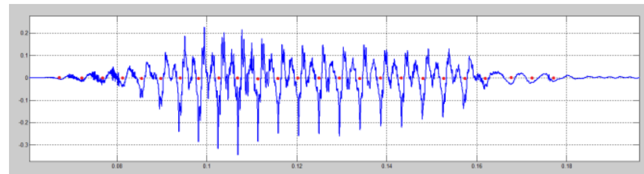

Fig.5.Pitch Detection model

1) Frequency-domain approaches:

Popular frequency domain algorithms include: the harmonic product spectrum: costrel analysis and maximum likelihood which attempts to match the frequency domain characteristics to pre-defined frequency maps to the detection of peaks due to harmonic series.

2) Spectral/temporal approaches:

Spectral/temporal pitch detection algorithms, pitch tracking are based upon a combination of time domain processing using an autocorrelation function such as normalized cross correlation, and frequency domain processing utilizing spectral information to identify the pitch.

3) Fundamental frequency of speech:

The fundamental frequency of speech can vary from 40 Hz for low-pitched male voices to 600 Hz for children or high-pitched female voices.

V. EXPERIMENTAL RESULTS

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Issue 3, March 2017

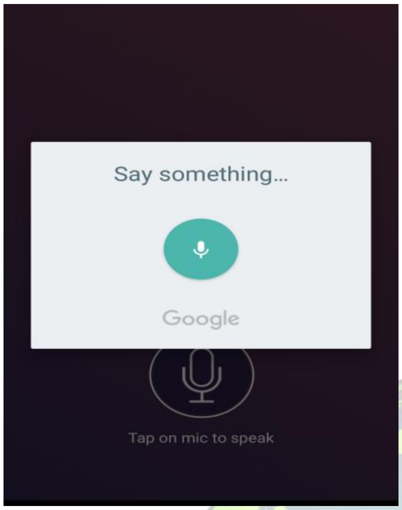

Fig.6.Waiting for user to give the input

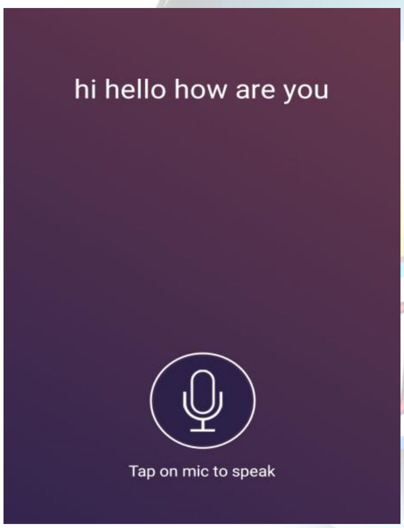

Fig.7.Input is displayed as text

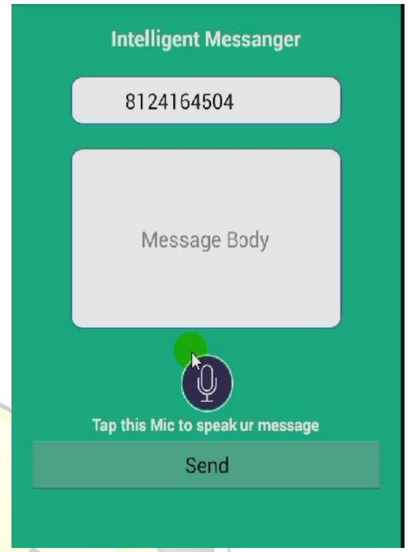

Fig.8.Application screenshot

VI. CONCLUSION

An automatic speech recognizer studied and implemented on the android platform which gives much accuracy for both numeric and alpha numeric inputs. The accuracy of this system is about 90%, and delay for recognition is less than 100 ns. We plan to implement this work for other languages as well as test them on the SMS sending application which is developed.

ACKNOWLEDGEMENT

We wish to express our gratitude to **Dr.N.KOTTISWARAN M.E., Ph.D.,** Principal, Info Institute of Engineering, Coimbatore, who always helped us whenever we approached him during the course of our project.

We would also like to express gratitude to **Dr.P.RAJKUMAR, M.E., Ph.D.,** Associate professor, Head of the Department, Department of computer science and Engineering, for his constant encouragement, guidance and constructive criticism in making this project a successful one.

We pay our respects and kindly thank our project co-ordinator **Mr.A.AROCKIA SELVARAJ M.E.,MBA.,** Assistant Professor, Department of Computer Science and

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Issue 3, March 2017

Engineering, for his valuable guidance and timely suggestion in bringing out this project.

We pay indebtedness to our project guide **Mr.B.JAYARAM, M.E.,** Assistant Professor, Department of Computer Science and Engineering, for his valuable guidance and timely suggestion in bringing out this project.

We express our thanks to all **faculty members** and **skilled assistants** of Computer Science and Engineering Department, my **family** and **friends** who helped me in every possible way.

REFERENCES

B. Raghavendhar Reddy, E. Mahender, "Speech to text conversion using android platform ", IJERA Feb-2013.

A.Mane, A.B.Patil," Comparison of LDM and HMM for an application of a speech", International Conference on Advances in Recent Technologies in Communication and Computing, 978-0-7695-4201- 0/10 $26.00 © 2010 IEEE, DOI 10.1109/ARTCom.2010.65-431

Rahim Patel, Dr. Y. SrinivasRao," Speech recognition using HMM with MFCC- an analysis using frequency spectral decomposition technique", SIPIJ Dec 2010.

Patrick Gampp, "Hidden markov model basics".

GulbaksheeDharmale, Dr. Vilas Thakare, Dr. Dipti D. Patil, "Intelligent HandsFree Speech based SMS System on Android" IEEE CONFERENCE 2016.

BIOGRAPHY

B.Jayaram is currently working as Assistant professor in Department of computer science and engineering at Info Institute of Engineering. He has obtained B.E from Bharathiar University in the year 2003 and M.E from Anna University Chennai in the year 2006. He has experience around 10 years and his area of interest include data mining, computer networks. He has guided many teams for doing projects. And he is a life time member of ISTE professional society.

GOKILAPRIYA.M is pursuing B.E, Computer Science Department at Info Institute of Engineering. She has an area of interest in Mobile Application Development.

THAMARAISELVI.R is pursuing B.E Computer Science and Engineering at INFO Institute of Engineering. Her area of interests are website designing and PHP.

KIRUTHIKA.M is currently pursuing B.E Computer Science and Engineering at INFO Institute of Engineering. Her area of interest lies in the field of Java, Android development and Software development.

MOHANA PRIYA.P is pursuing B.E Computer Science and Engineering at INFO Institute of Engineering. Her interest lies in Mobile Application Development and DBMS.