

A Review on Speech Interface for Form Filling Application

¹Priyanka Dahariya, ²Mrs. Shalu K Rakesh

¹M.Tech Scholar, ²Assistant Professor,

CSE Department

Chouksey Engineering College, Bilaspur, Chhattisgarh, India

Abstract: Speech recognition technology is one from the fast growing engineering technologies. Nearly 20% people of the world are suffering from various disabilities; many of them are blind or unable to use their hands effectively. They can share information with people by operating computer through voice input. The Speech is most prominent & primary mode of Communication among of human being. The communication among human computer interaction is called human computer interface. Speech has potential of being important mode of interaction with computer .It gives an overview of major technological perspective and appreciation of the fundamental progress of speech recognition and also gives overview technique developed in each stage of speech recognition. Our project is capable to recognize the speech and convert the input audio into text; it also enables a user to perform operations such as open calculator, word pad, notepad, log off computer. It concludes with the decision on feature direction for developing technique in human computer interface system using English Language.

Keywords: Speech recognition, HTML, Human computer interface

I INTRODUCTION

This report considers an overview of speech recognition technology, software development, and its applications. The first section deals with the description of speech recognition process, its applications in different sectors, its flaws and finally the future of technology. Later part of report covers the speech recognition process, and the code for the software and its working. Finally the report concludes at the different potentials uses of the application and further improvements and considerations.

Project Objective

To understand the speech recognition and its fundamentals. Its working and applications in different areas Its implementation as a desktop Application. Development for software that can mainly be used for :Speech Recognition, Speech Generation ,Text Editing ,Tool for operating Machine through voice.

Project Design

This project is designed and developed keeping that factor into mind, and a little effort is made to achieve this aim. Our project is capable to recognize the speech and convert the input audio into text; it also enables a user to perform operations such as “save, open, exit” a file by providing voice input. It also helps the user to open different system software such as opening Ms-paint, notepad and calculator. At the initial level effort is made to provide help for basic operations as discussed above, but the software can further be updated and enhanced in order to cover more operations

Project scope

This project has the speech recognizing and speech synthesizing capabilities though it is not a complete replacement of what we call a NOTEPAD but still a good text editor to be used through voice. This software also can open windows based software such as Notepad, Ms-paint and more.

Components of Speech recognition System

Voice Input

With the help of microphone audio is input to the system, the pc sound card produces the equivalent digital representation of received audio [8] [9] [10].

Digitization

The process of converting the analog signal into a digital form is known as digitization [8], it involves the both sampling and quantization processes. Sampling is converting a continuous signal into discrete signal, while the process of approximating a continuous range of values is known as quantization.

Acoustic Model

An acoustic model is created by taking audio recordings of speech, and their text transcriptions, and using software to create statistical representations of the sounds that make up each word. It is used by a speech recognition engine to recognize speech [8]. The software acoustic model breaks the words into the phonemes [10].

Speech Recognition Process

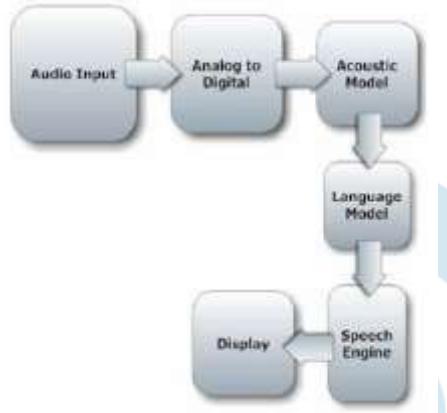


Figure 1.1 Speech Recognition Process

II LITERATURE REVIEW

H. Ghadage et.al (1) The current work presents a multilingual speech-to-text conversion system. Conversion is based on information in speech signal. Speech is the natural and most important form of communication for human being. Speech-To-Text (STT) system takes a human speech utterance as an input and requires a string of words as output.

Zhemin Tu et.al (2) speech recognition system based on an Internet client-server model is presented in this paper. A Java applet records the voice at the client computer, sends the recorded speech file over the Internet, and the server computer recognizes the speech and displays the recognized text back to the user. Using this structure, an isolated digit recognition application was realized.

R. Marin et.al (3) This article shows the way a speech recognition module has been implemented that provides the translation between voice input and the convenient text based commands to be executed on an already existing web robot. The novel contribution is the way the procedure is defined to be run over the Internet, and the interface implemented to connect any kind of external speech recognition program to the robot controller.

B. Raghavendar Reddy et.al (4) for the past several decades, designers have processed speech for a wide variety of applications ranging from mobile communications to automatic reading machines. Speech recognition reduces the overhead caused by alternate communication methods. Speech has not been used much in the field of electronics and computers due to the complexity and variety of speech signals and sounds.

Halimah B.Z. et.al (5) The Internet has became an important tool for learners to acquire information and knowledge that encompasses various elements such as text, graphic, numeric, and animation for their learning process.

Ryuichi Nisimura (6) we have developed a speech-to-text input method for web systems. The system is provided as a JavaScript library including an Ajax like mechanism based on a Java applet, CGI programs, and dynamic HTML documents. It allows users to access voice-enabled web pages without requiring special browsers. Web developers can embed it on their web page by inserting only one line in the header field of an HTML document. This study also aims at observing natural spoken interactions in personal environments. We have succeeded in collecting 4,003 inputs during a period of seven months via our public Japanese ASR server. In order to cover out-of-vocabulary words to cope with some proper nouns, a web page to register new words into the language model are developed. As a result, we could obtain an improvement of 0.8% in the recognition accuracy. With regard to the acoustical conditions, an SNR of 25.3 dB was observe

III PROBLEM STATEMENT

Speech Recognition Problem Definition: find accurate written transcription of spoken utterances.

Transcriptions may be in words, phonemes, syllables, or other units.

Accuracy: typically measured in terms of the edit distance between reference transcription and sequence output by the model.

Speaker verification.

Speaker identification.

Spoken-dialog systems.

Detection of voice features, e.g., gender, age, dialect, emotion, height, weight! Speech synthesis.

Highly variable: the same words pronounced by the same person in the same conditions typically lead to different waveforms.

IV PROPOSED METHODOLOGY

As an emerging technology, not all developers are familiar with speech recognition technology. While the basic functions of both speech synthesis and speech recognition takes only few minutes to understand (after all, most people learn to speak and listen by age two), there are subtle and powerful capabilities provided by computerized speech that developers will want to understand and utilize. Despite very substantial investment in speech technology research over the last 40 years, speech synthesis and speech recognition technologies still have significant limitations. Most importantly, speech technology does not always meet the high expectations of users familiar with natural human-to-human speech communication. Understanding the limitations - as well as the strengths - is important for effective use of speech input and output in a user interface and for understanding some of the advanced features of the Java Speech API.

An understanding of the capabilities and limitations of speech technology is also important for developers in making decisions about whether a particular application will benefit from the use of speech input and output

Speech Synthesis

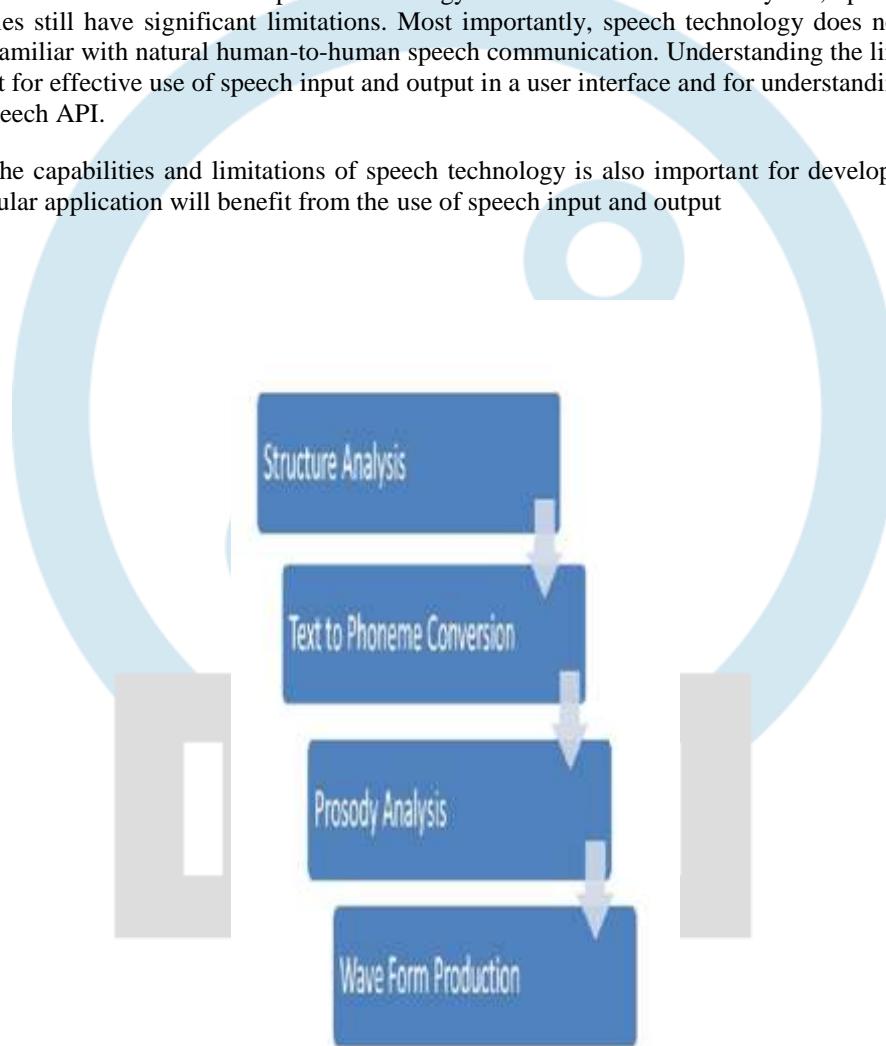


Figure 4.1 Speech Synthesis

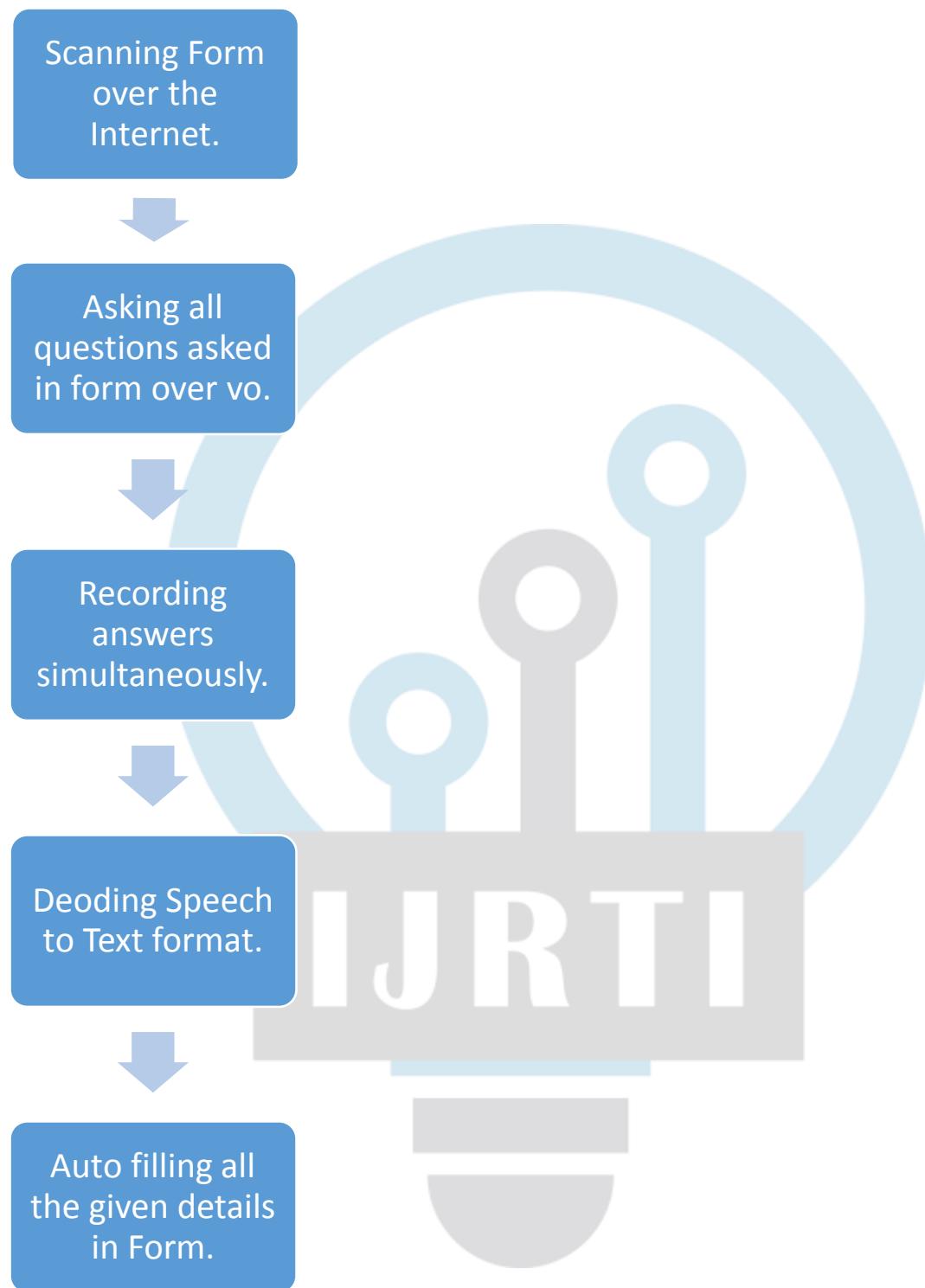
A speech synthesizer converts written text into spoken language. Speech synthesis is also referred to as *text-to-speech* (TTS) conversion. The major steps in producing speech from text are as follows:

Structure analysis: process the input text to determine where paragraphs, sentences and other structures start and end.

Synthesizer as an Engine

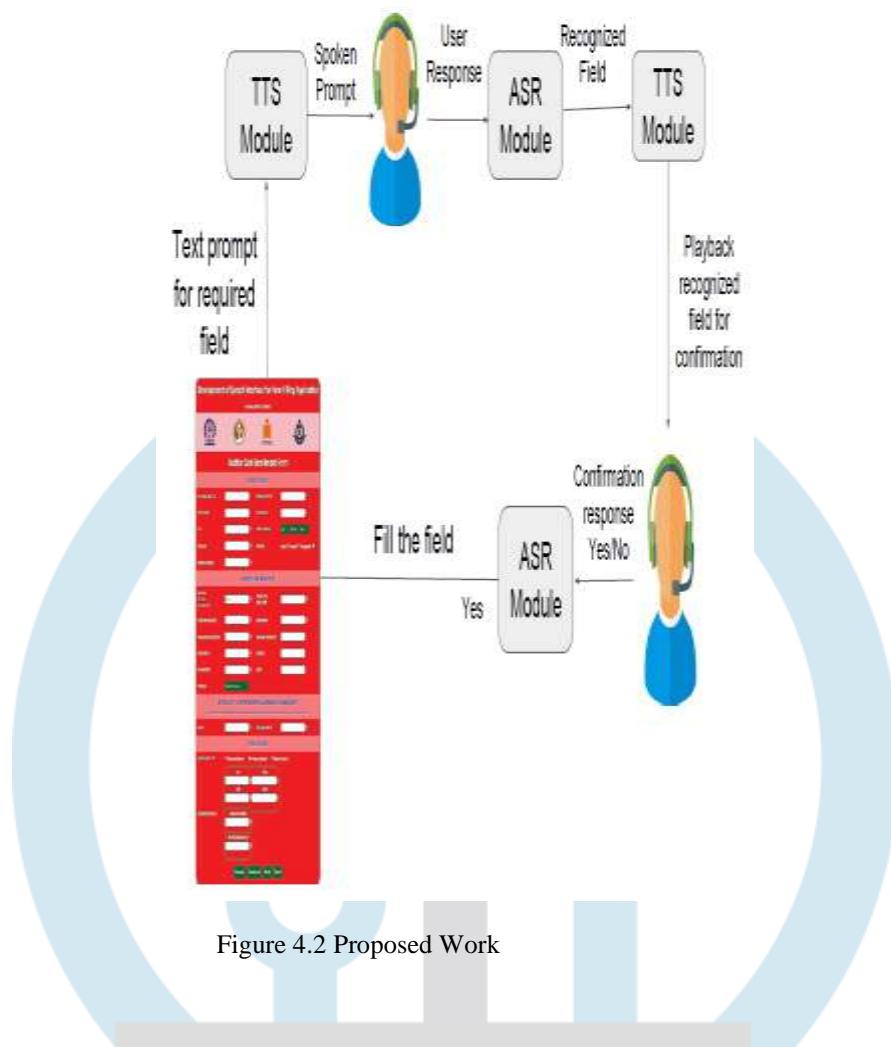
The basic functionality provided by a Synthesizer is speaking text, management of a queue of text to be spoken and producing events as these functions proceed. The Synthesizer interface extends the Engine interface to provide this functionality.

The following is a list of the functionality that the javax.speech. Synthesis package inherits from the javax.speech package and Figure 5.1 Interface for online form filling.



V EXPECTED RESULT

This Project work of speech recognition started with a brief introduction of the technology and its applications in different sectors. The project part of the Report was based on software development for speech recognition. At the later stage we discussed different tools for bringing that idea into practical work. The proposed work is developed. Finally it will be tested and results were discussed for the Speech Interface for Form Filling Application



REFERENCES

- [1] "Speech recognition- The next revolution" 5th edition.
- [2] Ksenia Shalonova, "Automatic Speech Recognition" 07 DEC 2007 Source:http://www.cs.bris.ac.uk/Teaching/Resources/COMS12303/lectures/Ksenia_Shalonova-Speech_Recognition.pdf
- [3] "Fundamentals of Speech Recognition". L. Rabiner & B. Juang. 1993. ISBN: 0130151572.
"Speech and Language Processing: An Introduction to Natural Language Processing, Computational
- [4]<http://www.abilityhub.com/speech/speech-description.htm>
- [5]<http://www.scribd.com/doc/2586608/speechrecognition.pdf> Date Added 04/21/2008
- [6]<http://electronics.howstuffworks.com/gadgets/high-tech-gadgets/speechrecognition.htm/printable> last updated: 30th oct 2009.
- [7]<http://www.jisc.ac.uk/media/documents/techwatch/ruchi.pdf>
- [8]Base paper was taken from the journal of Prerana Das, Kakalim Acharjee, Pranab Das and Vijay Prasad on SPEECH to TEXT Module.