ISSN: 0974-2115
www.jchps.com Journal of Chemical and Pharmaceutical Sciences

Speech and text recoginition for hearing and visually impaired people

HareeshBabu M.V*, Kalicharan.G, Ganesh Babu E.S
Jeppiaar Engineering College,Chennai
***Corresponding Author: E.mail:hareeshbabumv@gmail.com**

ABSTRACT

Many of the new technologies designed to help worldwide communication- e.g.: telephones, computers- have created new problems especially among the hearing and visually impaired. A person, who has severe hearing impairments, leads to deafness, may experience difficulties in communication. Conversely, someone with visual impairments has the similar difficulties. This paper discusses the issues and challenges met by visually and vocally challenged people. It also led to the development of language interpretation between them. Speech and text recognition is a technological implementation which is highly useful for hearing and visually impaired people to make their communication easier. Often people suffering from such disability have normal vision and therefore they are able to communicate through sign or gestures. This system through a wearable camera converts signs into text with the help of onboard electronic system. The other person who may not know sign language can read this text. Also the same text can be heard as speech by onboard voice synthesizer. Further, when the other person speaks the system will convert the speech into a text and sign language with the help of onboard voice recognizer system and communicator respectively. Hence this system makes communication easier between vocally and visually challenged people.

KEYWORDS: Onboard voice synthesizer, onboard electronic system, onboard voice recognizer, wearable camera, communicator, microphone, LCD.

INTRODUCTION

The application of speech recognition (SR) has improved over past several years; many of the proposed benefits have been demonstrated in varied applications. Speech recognition also referred to as speech-to-text or voice recognition, its technology that recognizes speech, allowing voice to serve as the "main interface between the visually and vocally impaired people".

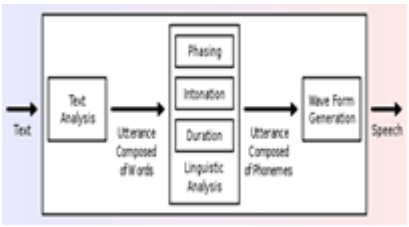
The goal of this presentation is to incorporate current speech recognition (speech-to-text) and speech synthesis (text-to-speech) technology, thus, providing solution to communication between the hearings bands visually impaired. Practically language interpretation between them is quite difficult. As, the sign made by the vocally challenged person cannot be visualized by a visually impaired person. In order to overcome this problem a new technological based system called "SPEECH AND TEXT RECOGNITION SYSTEM" has been found. This system converts the sign language to voice and the voice to sign. As an attempt we added a communicator to convert the text into video gestures, which further makes their communication better. Hope so this gives the overall view of our presentation.

Literature review: For most people, communication with others is quite simple. There are many options available: telephone, electronic-mail, instant messaging, etc. However, this becomes a more difficult task for those with disabilities. A deaf person does not have the luxury of being able to talk without additional equipment. Similarly, it is difficult for a blind person to communicate through electronic means that require being able to see the screen. Direct communication between a deaf and blind person is almost impossible without a mediator. There currently does not appear any software that directly addresses communication between the blind and the deaf. Much of the technology found is quite expensive, involves additional hardware.

Background: The main focus of this application is to provide an environment in which speech synthesis and recognition may be successfully implemented. This is done in the hopes that the accuracy ratings of this system would range in the 90-percentile. In order to accomplish, it was essential to develop strong background knowledge of the current processes and terminologies associated with speech synthesis and recognition.

Onboard electronic system: The onboard electronic system here we used through a wearable camera obtains text from sign. The other person who may not know sign language may read this text. The vocally disabled can keep his hand in front of the camera and make gestures using sign language. This camera captures the gestures and the onboard electronic system will identify and converts the gestures into text.

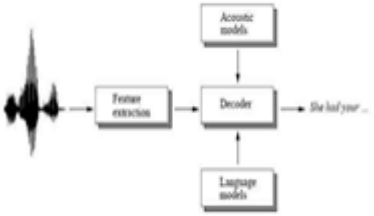
Speech synthesis: The speech synthesis involves two basic processes: the reading in of text and the production into sound. The device used for speech synthesis is called speech synthesizer. Let's discuss about the speech synthesizer below;



Specification: The SPE030 speech synthesizer module is complete text-to-speech 'talking' module, consists of on-board amplifier, speakers, and electronic circuit based upon the WTS701 single chip synthesizer (which also includes a text-to-speech processor).

Operation: A speech synthesizer is a device that undergoes three analysis text analysis, linguistic analysis, and wave form analysis. When a text input is given it gets analyzed in the text analyzer which results in utterance composed of words. Secondly, these utterances pass on to the linguistic analyzer where it gets modified in phasing intonation and duration techniques. The output utterance that composed of phonemes is generated as waveform in waveform generator. The resulting waveform gets converted into speech using a microphone.

Onboard voice recognizer: An onboard voice recognizer is a device which reverses the operation of a voice synthesizer (i.e.) it converts speech-to-text. It involves the following processes like speech analyzing, sound analyzing, and converting it into text.



Operation: The extraction of acoustic signal is the first step in speech recognition system. Now, the extracted acoustic signal pass on to the decoder which converts the signal into text. The decoder has other two inputs from acoustic models and language models. In acoustic model the frequency duration and nature of the signal is analyzed. Therefore, that output is given as one of the input for the decoder. To identify the language some language models are given as samples to the input of decoder. Thus, the decoder identifies the language along with the acoustic signals that converts the speech into text. The text is now displayed on the LCD screen. The acoustic signals are now converted to text.

Types of speech recognizer: There are two types of speech recognizer,

- Speaker - dependent
- Speaker – independent

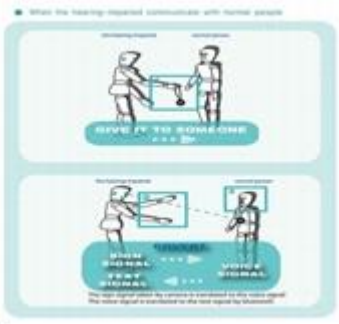
Speaker - dependent systems: Speaker-dependent recognition system is designed to be used by one person. To operate accurately, the system will need to be 'trained' to the user's individual speech patterns. This is sometimes referred to as "enrollment" of the speaker with the system. The speech patterns for the users will be recorded and pattern from which a template will be created for use by the speech recognizer. The major drawback of the system is that it is dedicated to a single user, and that it must be trained prior to its use. The speech template can be done at separate training stations prior to using the target system by transferring the created user voice templates to the target system. If more than one user is anticipated, or if the training of the system is not desirable, a speaker-independent system be an option.

Speaker – independent recognizers: Speaker-independent recognizer system is intended to allow multiple users to access a system using voice input. Examples of speaker-independent systems are directory accessed programs and an airline reservation system with a voice input driven menu system. Major drawbacks with the speaker-independent systems, in addition to increased complexity and difficult implementation, are its accuracy rate is low and slower response time. The impacts of these drawbacks continue to lessen with increased process of speeds, faster hardware and increased data storage capabilities. The speech signal may also include undesired noise signal. The process of noise filtering increases the efficiency by eliminating noise interference on corrupted speech signal regarding the acoustic quality in spoken word recording. Signal filter is most required for many tasks which can be considered into analog and digital filters. Filter is one of the most important parts of electronic circuit when it is designed to deal with signal processing directly and indirectly in some circumstances as well. The filter is designed to increase the stability, frequency range, low cost implementation and also increases efficiency in speech synthesis.

Liquid crystal display [LCD]: The screen used to display the text and the video gestures which makes communication easier.

Communicator: The software called communicator is used to convert the speech to video gestures which makes the communication much better.

Outlook of the process:



Speech and text recognition system processes sign language motions and can translate them using specific algorithms into an electronic voice. The prototype is based around a handheld device with a build-in microphone, speaker, sound board, video camera and a screen. The device can be a wearable one where it reads a user's sign language movements. Thus processes sign language motions, translating them into an electronic voice. Further it can also monitor a person's voice and translate words into sign language, projected on a LCD display. Here we have tried to translate the speech into the sign language. For this purpose we just added communicator in addition to the system's construction.

Advantages of the system:

- The device may be light weighted and therefore the user may wear this system on his body.
- It is a portable one.
- This system makes communication easier between vocally disabled and visually disabled people.

Challenges ahead:

- There is only one commercial software to convert speech to sign language which is very costlier.
- Translation can be done only in English.

Future research:

- Displaying both the text and video gestures on the LCD screen once the speech input is given, which makes the communication easier.
- Designing communicator at an affordable rate.
- Designing an onboard electronic system to facilitate varied language.

REFERENCES

A speech recognition and synthesis tool by Laura Schindler – at Stetson University Deland, Florida.

Speech recognition and synthesis, Douglas W Beeks.