

# SMART EMERGING TECHNOLOGY FOR BETTERMENT OF VISUALLY CHALLENGED PEOPLE'S

Mrs. Shashi Prabha Anan  
Department of computer science & Engineering  
MCAET, Ambedkar nagar  
Akabarpur, India

Dr. R. B. Gautam (Assistant professor)  
Department of Agriculture & Engineering  
MCAET, Ambedkar nagar  
Akabarpur, India

**Abstract**— Concept to connect and make our society completely digitalized, need to link Visual impairment people also can participate in using smart devices. This paper is proposed for the betterment of visually challenged people's. Aim to contribute in life of visually challenged people to be a part of smart emerging technology and to make their life quite easy. This type of smart technology will also encourage developers to build something more useful for visually impaired or illiterate people, who also deserve an equal standard in society. Mind reading computer system play important role for visually challenged society. These people find themselves technologically more challenged. This is true especially in the case of social networking, which these people cannot do without external help. With advancement in technology, presently visually challenged people are not able to use computers on their own mainly because keyboards are not user-friendly to them. In this methodology, trying to avoid use of keyboards and speech recognition models are used.

**Keywords**— Speech recognition, Visual impairment, Voice-Mail for blind, mobile computer devices, mobile technology, mobile assistive technology, handheld assistive technology etc.

## I. INTRODUCTION

Technology has the potential to enhance individuals' ability to participate fully in societal activities and to live independently. According to the official statistics from World Health Organization, there are about 285 million visually impaired persons in the world up to the year of 2011: about 39 million are completely blind and 246 million have weak sight [1]. This number will increase rapidly as the baby boomer generation ages [2]. These visually impaired people have great difficulty in perceiving and interacting with the surroundings, especially those which are unfamiliar. Fortunately, there are some navigation systems or tools available for visually impaired individuals. Traditionally, most people rely on the white cane for local navigation, constantly swaying it in front for obstacle detection [3]. However, they cannot adequately perceive all the necessary information such as volume or distance, etc. [4]. Since the weak sighted people have some degree of visual perception, and vision can provide more information than other senses, e.g. touch and hearing, the visual enhancement, which uses the popular AR (Augmented Reality) [10], [11] technique for displaying the surroundings and the feasible direction on the eyeglasses, is proposed to help the users to avoid the obstacle.

## II. HISTORY

The concept of speech recognition started somewhere in 1940s, practically the first speech recognition program was appeared in 1952 at the bell labs, that was about recognition of a digit in a noise free Environment. 1940s and 1950s consider as the foundational period of the speech recognition technology, in this period work was done on the foundational paradigms of the speech recognition that is automation and information theoretic models.

In the 1960's we were able to recognize small vocabularies (order of 10-100 words) of isolated words, based on simple acoustic-phonetic properties of speech sounds. The key technologies that were developed during this decade were, filter banks and time normalization methods. In 1970s the medium vocabularies (order of 100-1000 words) using simple template-based, pattern recognition methods were recognized. In 1980s large vocabularies (1000-unlimited) were used and speech recognition problems based on statistical, with a large range of networks for handling language structures were addressed. The key invention of this era were hidden markov model (HMM) and the stochastic language model, which together enabled powerful new methods for handling continuous speech recognition problem efficiently and with high performance.

In 1990s the key technologies developed during this period were the methods for stochastic language understanding, statistical learning of acoustic and language models, and the methods for implementation of large vocabulary speech understanding systems. After the five decades of research, the speech recognition technology

## III. RELATED WORK

This section we present a brief review of literature to highlight the previous related work in the field of speech interfaces. We start with a brief overview of ASR and TTS which form the backbone of such applications. Mostly systems working on speech Recognition are highly dependent on training done by the speaker. They only respond to the input given by the particular speaker making the system speaker dependent.

Different techniques have been used to enhance the system and to make its response more efficient and accurate i.e. Hidden Markov Model (HMM), Mel-Cepstral feature extraction algorithm and clustering techniques. Further adding to the recognition rate has developed a dictation pad for which they conducted an experiment in which they prove that age of the speaker also affects the recognition rate of the system. Young and old people have different pitch and frequencies which contribute a lot to the recognition part. For that purpose the training session was conducted with almost every age group of people. But now a day's recent speech APIs tools provide multi user recognition. We have trained our system in such a way that it will respond to the commands given by any user. And further will be able to perform operations based on that commands. Errors that encountered during the recognition of speech must be handled by some exception handling technique. Present assimilation tool which helps to execute applications in virtual environment. The purpose of this tool is to find speech recognition errors by performing consecutive simulations. Each and every action performed by the application will be considered as an event. The errors related to particular event will be displayed i.e. whether it is due to recognition problem or no specific event is linked to that action etc. Report a platform where applications can be developed for speech synthesis. The users don't have to install TTS systems on their machines rather than an online platform is provided to do text to speech synthesis.

#### IV. PROBLEM IDENTIFICATION

The existing system's does not provide a complete voice based application leading to lack of privacy and security for a visually impaired person. So, there emerged a need to create a complete voice based application for such people so that they can send or receive the emails with ease providing security and privacy. Email has turned out to be an important part of formal communication in professional world. For people who can see, emailing is not a big deal, but for people who are not blessed with ability of vision it poses a key concern because of its intersection with many vocational responsibilities. Thus, VMAIL turns out to be an application through which they can send or receive emails not only easily but securely too. For security purpose an additional feature of App Lock was included. App Lock is a security module which acts as an authenticator. It is the second layer of security which allows only genuine user to access the application. The screen was divided into 6 equal parts keeping in mind blind people's convenience. They had to save pass-code first time while logging in by shaking their phone. Then every time the user signs in to the system, he has to type the pass-code to get access to the application. Since our application is voice based thus our user doesn't need to remember the keypad layout of their android phones.

#### V. EXISTING SYSTEM

There are a total number of 4.1 billion email accounts created until 2014 and an there will be estimated 5.2 billion accounts by end of 2018[4] this makes emails the most used form of communication. The existing email systems don't provide any means of feedback or Talkback service. So, they cannot be used by visually challenged people. The voice based systems that have been developed till now are desktop applications. As visually challenged People do not prefer to have laptops and desktops, the use of such developed system are very limited. Even if the Braille keyboard is used, they have to remember the complex keyboard. The most common mail services that we use in our day today life cannot be used by visually challenged people. This is because they do not provide any facility so that the person in front can hear out the content of the screen. As they cannot visualize what is already present on screen they cannot make out where to click in order to perform the required operations. For a visually challenged person using a computer for the first time is not that convenient as it is for a normal user even though it is user friendly. Although there are many Screen readers available then also these people face some minor difficulties. Screen readers read out whatever content is there on the screen and to perform those actions the person will have to use keyboard shortcuts as mouse location cannot be traced by the screen readers. This means two things; one that the user cannot make use of mouse pointer as it is completely inconvenient if the pointer location cannot be traced and second that user should be well versed with the keyboard as to where each and every key is located. A user is new to computer can therefore not use this service as they are not aware of the key locations. Another drawback that sets in is that screen readers read out the Content in sequential manner and therefore user can make out the contents of the screen only if they are in basic HTML format. Thus the new advanced web pages which do not follow this paradigm in order to make the website more user-friendly only create extra hassles for these people. All these are some drawbacks of the current system which we will overcome in the system we are developing.

#### VI. PROPOSED SYSTEM

The new system will be designed in such a way that the user won't require to use the keyboard allot. The system will prompt the user with voice commands to perform certain action and the user will respond to the same. The main benefit of this system is that the use of keyboard is completely eliminated. The user will have to respond through voice and mouse click only.

The proposed system is based on a completely novel idea and is nowhere like the existing mail systems. The most important aspect that has been kept in mind while developing the proposed system is accessibility. A web system is said to be perfectly accessible only if it can be used efficiently by all types of people whether able or disable. The current systems do not provide this accessibility. Thus the system we are developing is completely different from the current system. Unlike current system which emphasizes more on user friendliness of normal users, our system focuses more on user friendliness of all types of people including normal people visually impaired people as well as illiterate people.

The complete system is based on IVR- interactive voice response. When using this system the computer will be prompting the user to perform specific operations to avail respective services and if the user needs to access the respective services then he/she needs to perform that operation. One of the major advantages of this system is that user won't require to use the keyboard. All operations will be based on mouse click events. Now the question that arises is that how will the blind users find location of the mouse Pointer. The system has given the user a free will to click blandly anywhere on the screen. Which type of click will perform which function will be specified by the IVR. Thus user need not worry about location of the mouse at all. This system will be perfectly accessible to all types of users as it is just based on simple mouse clicks and speech inputs and there is no need to remember keyboard shortcuts. Also because of IVR facility those who cannot read need not worry as they can listen to the prompting done by the system and perform respective actions.

The system consists of following components as depicted in above figure:

##### 1. VMAIL Android App

The basic function of the VMAIL application is to provide user with a simple way to perform email operations on his phone without compromising his security. The application is totally voice-based allowing blind person to send and receive emails on the go. It converts the user spoken voice into text and performs the action accordingly. It consists of voice confirmation i.e., confirming if the user has actually spoken the

recognized text or not, which minimizes the errors involved. It has the feature of address book which helps user to save the email as a contact and use that contact in future for sending the mail.



Fig.6.1 VMAIL demo

## 2. App Lock

App Lock is a security module which acts as an authenticator. It is second layer of security which allows only genuine user to access the application. The screen was divided into six equal parts keeping in part blind people's convenience. They had to save pass code first time while logging in by shaking their phone. Then every time the user signs in to the system, he has to type the pass code to get access to the application.



Fig 6.2. App Lock demo

## 3. SQLite

SQLite is a software library that implements a self-contained, server-less, zero-configuration, transactional SQL database engine. It is the default database provided by Android. All the address book contacts along with email addresses resides on user's phone in SQLite database.

## 4. Java Mail API

The Java Mail API provides a platform-independent and protocol-independent framework to build mail and messaging applications. The Java Mail API provides a set of abstract classes defining objects that comprise a mail system. It is an optional package (standard extension) for reading, composing, and sending electronic messages.

## 5. Backend

The VMAIL application is connected to the email provider using internet. The emails are sent using SMTP protocol, whereas, they are retrieved from the server using IMAP protocol. The Java Mail API interacts with the Google's SMTP and IMAP server through the internet to send and retrieve the emails. The communication is secured using HTTPS and SSL protocols.



Fig.6.3. VMAIL Speech Recognition

## VII. 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. 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:

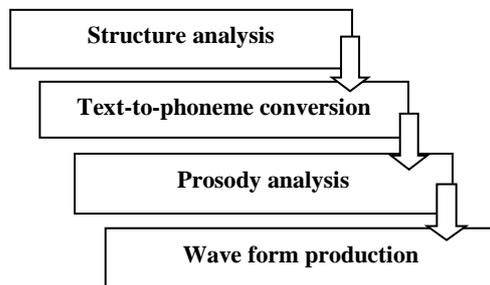


Fig.6.4. Speech Synthesis

- **Structure analysis:** process the input text to determine where paragraphs, sentences and other structures start and end. For most languages, punctuation and formatting data are used in this stage.
- **Text pre-processing:** analyze the input text for special constructs of the language. In English, special treatment is required for abbreviations, acronyms, dates, times, numbers, currency amounts, email addresses and many other forms.
- Other languages need special processing for these forms and most languages have other specialized requirements.
- The remaining steps convert the spoken text to speech.
- **Text-to-phoneme conversion:** convert each word phonemes. A phoneme is a basic unit of sound in a language. US English has around 45 phonemes including the consonant and vowel sounds. For example, "times" is spoken as four phonemes "t ay m s". Different languages have different sets of sounds (different phonemes).
- **Prosody analysis:** process the sentence structure, words and phonemes to determine appropriate prosody for the sentence. Prosody includes many of the features of speech other than the sounds of the words being spoken. This includes the pitch (or melody), the timing (or rhythm), the pausing, the speaking rate, the emphasis on words and many other features. Correct prosody is important for making speech sound right and for correctly conveying the meaning of a sentence.

**Wave form production:** finally, the phonemes and prosody information aroused to produce the audio waveform for each sentence. There are many ways in which the speech can be produced from the phoneme and prosody information. Most current systems do it in one of two ways: concatenation of chunks of recorded human speech, or formant synthesis using signal processing techniques based on knowledge of how phonemes sound and how prosody affects those phonemes. The details of waveform generation are not typically important to application developers.

## VII. CONCLUSION

Voice based architecture helps blind people to access e-mail with no difficulty. The proposed system entirely focuses on the benefit of the blind in making use of advanced technology for their growth and improvement. This design will also reduce cognitive load taken by blind to remember and type characters using keyboard. It also helps handicapped and illiterate people. This project will be very much useful for today's generation either blind or physically challenged to move a step forward in their way in an easy manner to achieve their desire. We have proposed a system which will help the visually impaired people to access email services efficiently. This system will help in overcoming some drawbacks that were earlier faced by the blind people in accessing emails. We have eliminated the concept of using keyboard shortcuts along with screen readers which will help reducing the cognitive load of remembering keyboard shortcuts. Also any naive user who does not know the location of keys on the keyboard need not worry as Keyboard usage is eliminated. The user only needs to follow the instructions given by the IVR and use mouse clicks accordingly to get the respective services offered. Other than this the user might need to feed in information through voice inputs when specified.

## REFERENCES

- [1] L. Tian, Y. Tian and C. Yi, "Detecting good quality frames in videos captured by a wearable camera for blind navigation," in 2013 IEEE Int. Conf. Bioinformatics and Biomedicine, Shanghai, 2013, pp. 334-337.
- [2] M. Moreno, S. Shahrabadi, J. José, J. M. H. du Buf and J. M. F. Rodrigues, "Realtime local navigation for the blind: Detection of lateral doors and sound interface," in Proc. 4th Int. Conf. Software Development for Enhancing Accessibility and Fighting Info-exclusion, 2012, pp. 74-82.
- [3] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: A survey," IEEE Trans. Systems, Man, Cybern., vol. 40, no. 1, pp. 25-35, Jan. 2010.
- [4] A. Aladrén, G. López-Nicolás, L. Puig and J. J. Guerrero, "Navigation assistance for the visually impaired using RGB-D sensor with range expansion," IEEE Systems J., vol. 10, no. 3, pp. 922-932, Sept. 2016,

- [5] H. Fernandes, P. Costa, V. Filipe, L. Hadjileontiadis and J. Barroso, "Stereo vision in blind navigation assistance," 2010 World Automation Congr., Kobe, 2010, pp. 1-6.
- [6] W. Heuten, N. Henze, S. Boll and M. Pielot, "Tactile wayfinder: a non-visual support system for wayfinding," Nordic Conf. Human Computer Interaction, Lund, 2008, pp. 172-181.
- [7] M. C. Kang, S. H. Chae, J. Y. Sun, J. W. Yoo and S. J. Ko, "A novel obstacle detection method based on deformable grid for the visually impaired," IEEE Trans. Consumer Electron., vol. 61, no. 3, pp. 376-383, Aug. 2015.
- [8] J. Sánchez, M. Sáenz, A. Pascual-Leone and L. Merabet, "Navigation for the blind through audio-based virtual environments," in Proc. 28th Int. Conf. Human Factors in Computing Syst., Atlanta, Georgia, 2010, pp. 3409-3414.
- [9] J. Kim and H. Jun, "Vision-based location positioning using augmented reality for indoor navigation," IEEE Trans. Consumer Electron., vol. 54, no. 3, pp. 954-962, Aug. 2008.
- [10] R.EL Kaliouby and P.Robinson- Automated Interference of cognitive mental states from Video, Computer laboratory University of Cambridge UK.
- [11] Jag tap Nilesh, Pawan Alai, Chavhan Swapnil and Bendre M.R. "Voice Based System in Desktop and Mobile Devices for Blind People". In International Journal of Emerging Technology and Advanced Engineering (IJETA), 2014 on Pages 404-407 (Volume 4, issue 2).
- [12] Ummuhanyifa U., Nizar Banu P K , "Voice Based Search Engine and Web page Reader". In International Journal of Computational Engineering Research (IJCER). Pages 1-5.
- [13] G. Shoba, G. Anusha, V. Jeevitha, R. Shanmathi. "AN Interactive Email for Visually Impaired". In International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), 2014 on Pages 5089-5092.(Volume 3, Issue 1).
- [14] World Health Organisation (WHO), 2011. Visual Impairment and Blindness: Fact Sheet No 282.

