

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

An Improved Text detection based on Morphological Operations and Inpainting

Rohini.P, Ragini.A, Ramya.B

Associate Professors, Department of Computer Engineering, Mount Zion College of Engineering and Technology,
Pudukkottai, India

ABSTRACT: Dictation using speech recognition input method for mailing devices for blind. This Application is used for Visually Challenged people to access mail efficiently. 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. Thus reliance of visually. The users of this system would not need to have any basic information regarding keyboard shortcuts or where the keys are located. All functions are based on simple mouse click operations making it very easy for any type of user to use this system Impaired on other people for their activities related to mail can be reduced. This application proposes an android application, designed specifically for visually impaired people. This application provides a voice based mailing service where they could read and send mail on their own, without any guidance. The users have to use certain keywords which will perform certain actions for e.g. Read, Send, Compose Mail, Address Book etc. The VMAIL system can be used by a blind person to access mails easily and efficiently This makes internet a completely useless technology for the visually impaired and illiterate people. Even the systems that are available currently like the screen readers TTS and ASR do not provide full efficiency to the blind people so as to use the internet. As nearly 285 million people worldwide are estimated visually impaired it become necessary to make internet facilities for communication usable for them also. Therefore we have come up with this project in which we will be developing a voice based email system which will aid the visually impaired people who are naive to computer systems to use email facilities in a hassle free manner. The users of this system would not need to have any basic information regarding keyboard shortcuts or where the keys are located. All functions are based on simple mouse click operations making it very easy for any type of user to use this system.

KEYWORDS: Text detection, Inpainting, Morphological operations, Connected component labelling.

I. INTRODUCTION

Internet is considered as a major store house of information in today's world. No single work can be done without the help of it. It has even become one of the de facto methods used in communication. And out of all methods available email is one of the most common forms of communication especially in the business world. However not all people can use the internet. This is because in order to access the internet you would need to know what is written on the screen. If that is not visible it is of no use. This makes internet a completely useless technology for the visually impaired and illiterate people. Even the systems that are available currently like the screen readers TTS and ASR do not provide full efficiency to the blind people so as to use the internet. As nearly 285 million people worldwide are estimated visually impaired it become necessary to make internet facilities for communication usable for them also. Therefore we have come up with this project in which we will be developing a voice based email system which will aid the visually impaired people who are naive to computer systems to use email facilities in a hassle free manner. The users of this system would not need to have any basic information regarding keyboard shortcuts or where the keys are located. All functions are based on simple mouse click operations making it very easy for any type of user to use this system. Also the user need not worry about remembering which mouse click operation he/she needs to perform in order to avail a given service as the system itself will be prompting them as to which click will provide them with what operations.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

II. WORK RECOGNITION

An easy way to Voice recognition software (also known as speech to text software) allows an individual to use their voice instead of typing on a keyboard. Voice recognition may be used to dictate text into the computer or to give commands to the computer. Voice recognition software allows for a quick method of writing onto a computer. It is also useful for people with disabilities who find it difficult to use the keyboard. This software can also assist those who have difficulty with transferring ideas onto paper as it helps take the focus out of the mechanics of writing. Word recognition is measured as a matter of speed, such that a word with a high level of recognition is read faster than a novel one. This manner of testing suggests that comprehension of the meaning of the words being read is not required, but rather the ability to recognize them in a way that allows proper pronunciation. Therefore, context is unimportant, and word recognition is often assessed with words presented in isolation in formats such as flash cards. Nevertheless, ease in word recognition, as in fluency, enables proficiency that fosters comprehension of the text being read.

III. SPEECH RECOGNITION

Speech recognition for application Voice SMS is done on Google server, using the HMM algorithm. HMM algorithm is briefly described in this part. Process involves the conversion of acoustic speech into a set of words and is performed by software component. Accuracy of speech recognition systems differ in vocabulary size and confusability, speaker dependence vs. independence, modality of speech (isolated, discontinuous, or continuous speech, read or spontaneous speech), task and language constraint. Speech recognition system can be divided into several blocks: feature extraction, acoustic models database which is built based on the training data, dictionary, language model and the speech recognition algorithm. Analog speech signal must first be sampled on time and amplitude axes, or digitized. Samples of speech signal are analyzed in even intervals. This period is usually 20 ms because signal in this interval is considered stationary. Speech feature extraction involves the formation of equally spaced discrete vectors of speech characteristics. Feature vectors from training database are used to estimate the parameters of acoustic models. Acoustic model describes properties of the basic elements that can be recognized. The basic element can be a phoneme for continuous speech or word for isolated words recognition. Dictionary is used to connect acoustic models with vocabulary words. Language model reduces the number of acceptable word combinations based on the rules of language and statistical information from different texts. Speech recognition systems, based on hidden Markov models are today most widely applied in modern technologies. They use the word or phoneme as a unit for modeling. The model output is hidden probabilistic functions of state and can't be deterministically specified. State sequence through model is not exactly known. Speech recognition systems generally assume that the speech signal is a realization of some message encoded as a sequence of one or more symbols. To effect the reverse operation of recognizing the underlying symbol sequence given a spoken utterance, the continuous speech waveform is first converted to a sequence of equally spaced discrete parameter vectors. Vectors of speech characteristics consist mostly of MFC (Mel Frequency Cepstral) coefficients, standardized by the European Telecommunications Standards Institute for speech recognition. The European Telecommunications Standards Institute in the early 2000s defined a standardized MFCC algorithm to be used in mobile phones. Standard MFC coefficients are constructed in a few simple steps. A short-time Fourier analysis of the speech signal using a finite-duration window (typically 20ms) is performed and the power spectrum is computed. Then, variable bandwidth triangular filters are placed along the perceptually motivated frequency scale and filter bank energies are calculated from the power spectrum. Magnitude compression is employed using the logarithmic function. Finally, auditory spectrum thus obtained is decorrelated using the DCT and first (typically 13) coefficients represent the MFCCs.

IV. OBJECTIVES

- [1] This application proposes an android application, designed specifically for visually impaired people
- [2] This application provides a voice based mailing service where they could read and send mail on their own, without any guidance.
- [3] The users have to use certain keywords which will perform certain actions for e.g. Read, Send, Compose Mail, Address Book etc.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

[4] The VMAIL system can be used by a blind person to access mails easily and efficiently.

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 this makes emails the most used form of communication. The most common mail services that we use in our day to day 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 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. As particular location cannot be tracked by the blind user 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

A. SPEECH_TO_TEXT CONVERTER

The system acquires speech at run time through a microphone and processes the sampled speech to recognize the uttered text. The recognized text can be stored in a file. We are developing this on Android platform using Eclipse workbench. Our speech-to-text system directly acquires and converts speech to text. It can supplement other larger systems, giving users a different choice for data entry. A speech-to-text system can also improve system accessibility by providing data entry options for blind, deaf, or physically handicapped users. Speech recognition system can be divided into several blocks: feature extraction, acoustic models database which is built based on the training data, dictionary, language model and the speech recognition algorithm. Analog speech signal must first be sampled at time and amplitude axes, or digitized. Samples of the speech signal are analyzed in even intervals. This period is usually 20 ms because the signal in this interval is considered stationary. Speech feature extraction involves

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

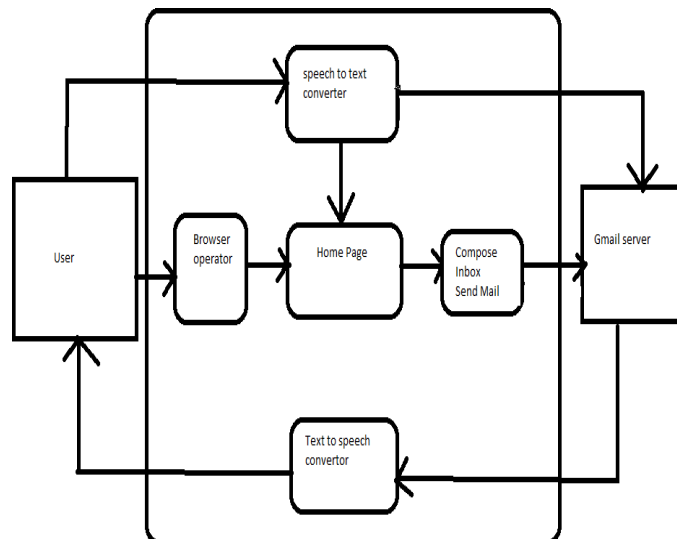
Visit: www.ijirset.com

Vol. 8, Issue 7, July 2019

the formation of equally spaced discrete vectors of speech characteristics. Feature vectors from training database are used to estimate the parameters of acoustic models. The acoustic model describes properties of the basic elements that can be recognized. The basic element can be a phoneme for continuous speech or word for isolated words recognition.

B. TEXT TO SPEECH

Converter should Converting text to voice output using speech synthesis techniques. Although initially used by the blind to listen to written material, it is now used extensively to convey financial data, e-mail messages, and other information via telephone for everyone. Text- to-speech is also used on handheld devices such as portable GPS units to announce street names when giving directions. Our Text-to-Speech Converterl accepts a string of 50 characters of text (alphabets and/or numbers) as input. In this, have interfaced the keyboard with the controller and defined all the alphabets as well as digits keys on it. The speech processor has an unlimited dictionary and can speak out almost any text provided at the input most of the times. Hence, it has an accuracy of above 90%. It is a microcontroller based hardware coded in Embedded C language. Further research is to be done to optimize various methods of inputting the text i.e. Reading the text using optical sensor and converting it to speech so that almost all sorts of physical challenges faced by the people while communicating are overcome



International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 6, June2019

VII. CONCLUSIONS

Thus conclude that 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.

VIII. FUTURE WORK

In future, For people who can see, e-mailing is not a big deal, but for people who are not blessed with gift of vision it postures a key concern because of its intersection with many vocational responsibilities. This voice based email system has great application as it is used by blind people as they can understand where they are. E.g. whenever cursor moves to any icon on the website say Register it will sound like "Register Button". There are many screen readers available. But people had to remember mouse clicks. Rather, this project will reduce this problem as mouse pointer would read out where he/she lies. This system focuses more on user friendliness of all types of persons including regular persons, visually compromised people as well as illiterate.

IX. RESULT

Algorithm	Accuracy (%)
Back propagation	10
Hidden markov model	20
Speech to text	70
Text to speech	80

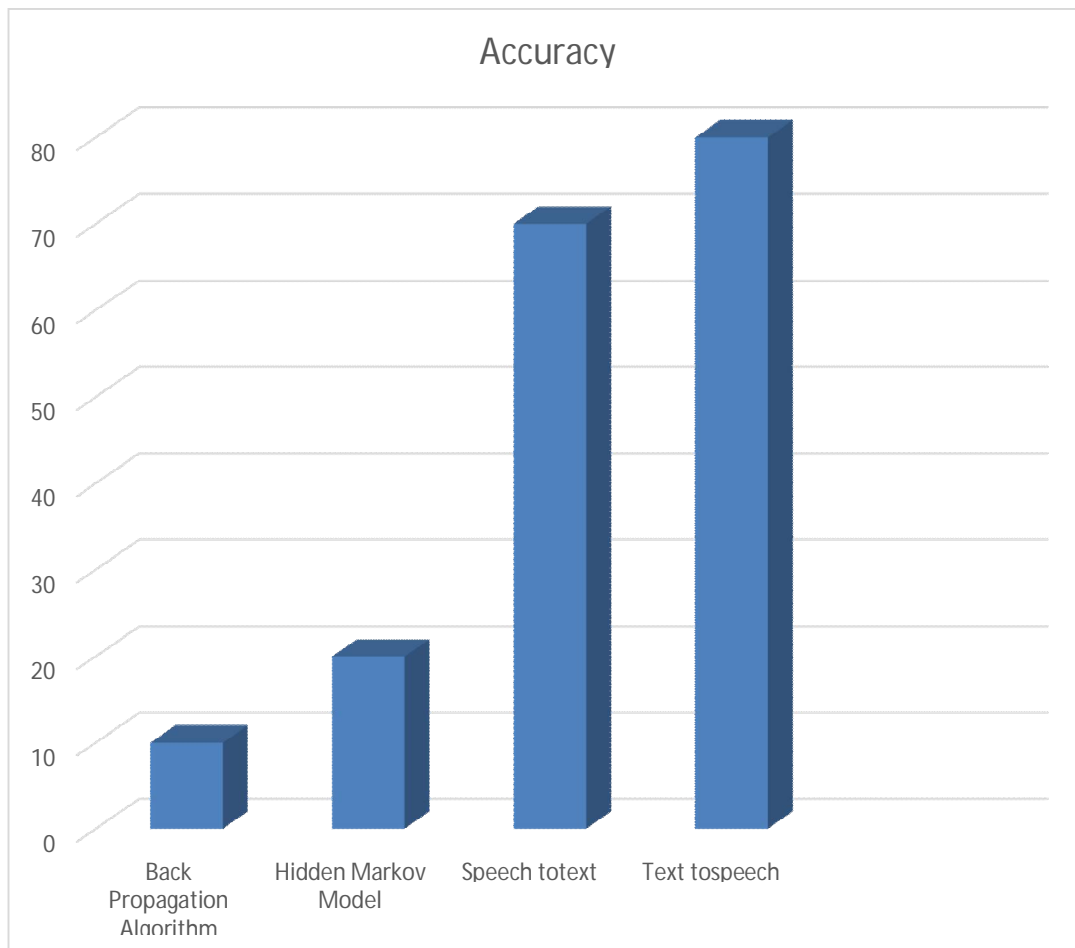
ACCURACY TABLE

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 6, June2019



GRAPH – ACCURACY

REFERENCES

- [1] C. Kang, H. Jo and B. Kim, "A Machine-to-Machine based Intelligent Walking Assistance System for Visually Impaired Person", The Journal of KICS, vol. 36, no. 3, (2011), pp.195-304.
- [2] S. Kumar, M. A. Qadeer and A. Aupta, "Location Based Service using Android", Internet Multimedia Service Architecture and Applications, IEEE International Conference,(2009).
- [3] H. -W. Jung, "Smartphones and future changes", The Korea Contents Association, vol. 8, no. 2,(2010).
- [4] I -H. O, J. S. Bae, D. -W. Park and Y. -H. Sohn, "Implementation of Location Based Service(LBS) using GPS for Various Sizes of Maps", Korean Institute of Information Technology, vol. 8, no. 4,(2010).
- [5] H. E. Lee and J. W. Lee, "Google Android phone Personal open market", Korean Multimedia Society, Fall Conference, (2009), pp. 346-349.
- [6] G. Broll, S. Keck, P. Holleis and A. Butz, "Improving the Accessibility of NFC/RFID- based Mobile Interaction through Learnability and Guidance", International Conference on Human-Computer Interaction with Mobile devices and services, vol. 11,(2009).
- [7] CheePeng Lim, Siew Chan Woo, AunSimLoh, Rohaizan Osman, "Speech Recognition Using Artificial Neural Networks," wise, vol. 1, pp.0419, First International Conference on Web Information Systems Engineering (WISE'00)-Volume 1,2000.
- [8] B.-s. Lin, H.-m. Wang and L.-s. Lee, "A distributed architecture for cooperative spoken dialogue agents with coherent dialogue state and history,"1999.
- [9] R. Privat, N. Vigouroux, C. Bousquet, P. Truillet and B. Oriola, "Speech technology for universal access in interactive systems?," in HCI,2001.
- [10] C. Callaway and K. Sima'an, "Wired for speech: How voice activates and Advances the human-computer relationship," Computational Linguistics, vol. 32, no. 3, pp. 451-452, 2006