

# Breaking New Grounds: Realization Of “ALL EAR” Screen Readers

Dr. M. A. Pradhan, Shambhavi Sinha, Swarali Shirke

Vedant Deshmukh, Akhilesh Mishra

Department of Computer Science and Engineering

AISSMS COE

Pune - India

## ABSTRACT

In the world, where people are relying more and more on smart devices for their work and existence, people who are deprived of the gift of vision are lacking behind in all the fields. Though with the invention of screen readers, many visually impaired people have regained their confidence but there are still many drawbacks of this technology. Here in our paper, we have presented the review on the current scenario of the present screen readers, their working challenges and the requirements of the blind people. We aim to provide a better understanding of the areas that are still under ignorance which can be crucial for future development.

**Keywords:-** Assistive Technology, Screen Readers, Text-to-speech, Job Access With Speech, Non Visual Desktop Access.

## I. INTRODUCTION

According to the WHO (World Health Organization) about 285 million people worldwide are visually impaired and of these, 39 million are blind. 90% of blind people live in low income countries. [1] Since the technology has become a major part of our lives, it is very difficult for the visually impaired people to keep up with their daily chore independently. Many inventions have been made in order to make the life of these people easy and to help them cope up with the fast growing industry by providing them assistant tools that help them fetch jobs and earn their living.

Speech is the most natural way of communication, it also provides an efficient means of man-machine communication. Speech interfacing provides a way to communicate with the machine without using keyboards or mouse. Speech interfacing involves speech recognition and speech synthesis. Speech recognition refers to the ability to listen spoken words and identify various sounds present in it and recognize them as words of some known language. Speech synthesizer takes the text as input and converts it into speech output that is it acts as text to speech converter [A]

Assistive technology is defined as any item, piece of equipment or product system that is used to increase, maintain or improve functional capabilities of individuals with disabilities. It is a generic term used for a wide range of software and hardware designed for disabled people to assist

with everyday tasks in both work and home life. [2] Many devices like screen readers, screen magnifiers, braille e-book are available for as the assistive tools. One of the important assistive tool (AT) is the screen readers. [3]

Special blind computers that has been developed play an important role in making these people self-reliant. These computers rely on an assistive technology that interprets and announces screen content via voice synthesizer, called screen readers. Some screen readers make use of crystal that can expand when exposed to particular voltage level (piezo effect) allowing visually impaired to use their fingers to read text.

Screen Reader can also be used in a stand-alone computer environment, as a terminal connected to a host, or in a network. It runs simultaneously with most application programs, enabling the user to hear complete screens, paragraphs, windows, highlighted areas, sentences, lines or characters. Software prefers tailored for specific applications such as data base managers, spreadsheets, word processors, and other productivity tools are included in the product. [4]

The screen reader mainly focuses on the audio output mechanism- text to speech (TTS) engine, which is primary output for a majority of people with vision impairments. The TTS refers speech synthesis, that is, an artificial production of human speech. A TTS system converts normal language text into speech and is generally independent of the screen reader

and building a high quality TTS can be a fairly effort-intensive project.

The main issue with the screen readers are the individual preference of the users and the ability of the screen reader to complete a specific task. The other problem with the screen readers is the inability to read text when the alternative text for images is not present. This area has always been compromised because of its complexity. A person’s choice of a screen reader depends on many factors like cost, operating system, the functionality supported by the reader, etc. Many companies like IBM, Microsoft, Apple have been making screen readers from past few decades. Some of the examples of the screen readers are Job Access with speech (JAWS), Non-Visual Desktop Access (NVDA), voice-over, windows-eye, etc. [6]

## II. LITERATURE SURVEY

In the past decade, a number of works have been done in the field of speech synthesis and speech recognition for communication. Many assistive tools and technology have been implemented to help the blind people to use recent technologies in their day-to-day life.

The study shown by the authors [7] describe a novel and intelligent tool called Swapn for visually challenged and aged people to use multimedia technologies by overcoming their disabilities. Multimedia technologies like image recognition, social networking and warning systems were used by using multimedia tools like image sensor, proximity sensor and camera laser sensor to model a Human Computer Interaction (HCI) system.

Another study [8] focused on the blind as well as physically handicapped people. In this study, the authors have presented a scheme to develop a voice recognition system which can be used to control a computer through user voice input without the use of mouse or keyboard. This voice operated computer called as Swar supports features like voice dialing, call routing, audio search and data entry through speech-to-text processing.

The authors [9] describe Homer, a voice based text-to-speech system for the blind or visually impaired people, to read and understand the Slovenian texts. In this, special corpora is organized on the computer network server from where the users can obtain texts. User can also control the system by giving some particular commands in the form of user voice input.

The study shown by the authors [10] describes the development of an assistive technology to help two physically handicapped persons, deaf and blind, to interact or communicate with each other in noise free environment. This technology is based on Speech Recognition and Synthesis Tool (SRST) which uses Microsoft Windows Speech Application Programming Interface (SAPI).

The researchers developed a new and trained system named as Drishti that works in both outdoor and indoor environment. It is a navigation system for visually impaired people. It uses DGPS as its location system for outdoor navigation. It provides optimal route by the means of dynamic routing and rerouting ability. For indoor location measurement, OEM ultrasound positioning system is used. Vocal prompts about the obstacles are given to the user and step-by-step walkthrough guidance is also provided [11].

This study [12] describes a voice based system on desktop and mobile devices which helps blind people to access e-mail and multimedia features of the operating system. The system includes features like Gmail, RSS (Real Simple Syndication), Drive browser, songs and book reader.

## III. PROPOSED SYSTEM

The proposed system is based on a novel idea. The most important aspect which has been kept in mind while developing the proposed is the accessibility. We aim to develop a voice recognition system which would be used for controlling computer via speech input from user without the use of mouse and keyboard. This system would be a great benefit to visually handicapped people. The basic layout of our system is depicted in the diagram below:-

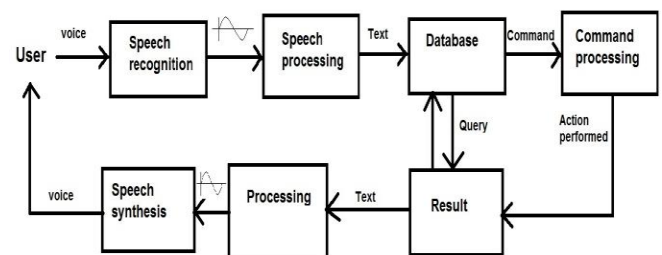


Fig. 1 Block diagram of the system

The key features of this system will be robustness, operability, perceivable and understandable. We have eliminated the use of keyboards and mouse by using vocal commands for better navigation. This also reduces the

problem of remembering the short cut or having to learn braille. The system is divided into the following major tasks:-

1. Speech Recognition
2. Speech Processing
3. Database handling
4. Speech Synthesis

The user gives the vocal commands as input to the system and the system works accordingly. No keyboard keys or mouse click is required. The database contains commands mapped to the actions that are needed to be performed. After the system analysis and produces result, it conveys the result in voice format to the user.

Speech recognition is the inter-disciplinary sub-field of computational linguistics which incorporates knowledge and research in the linguistics, computer science, and electrical engineering fields to develop methodologies and technologies that enables the recognition and translation of spoken language into text by computers and computerized devices such as those categorized as Smart Technologies and robotics. It is also known as "automatic speech recognition" (ASR), "computer speech recognition", or just "speech to text" (STT).

We are using SAPI (speech application programming interface) for speech processing. The Speech Application Programming Interface or SAPI is an API developed by Microsoft to allow the use of speech recognition and speech synthesis within Windows applications. To date, a number of versions of the API have been released, which have shipped either as part of a Speech SDK, or as part of the Windows OS itself. Applications that use SAPI include Microsoft Office, Microsoft Agent and Microsoft Speech Server.

In general all versions of the API have been designed such that a software developer can write an application to perform speech recognition and synthesis by using a standard set of interfaces, accessible from a variety of programming languages. In addition, it is possible for a 3rd-party company to produce their own Speech Recognition and Text-To-Speech engines or adapt existing engines to work with SAPI. In principle, as long as these engines conform to the defined interfaces they can be used instead of the Microsoft-supplied engines.

This is effective in windows XP operating system. Hence, we have used windows XP. The GUI has been created using Netbeans 7.1(IDE).

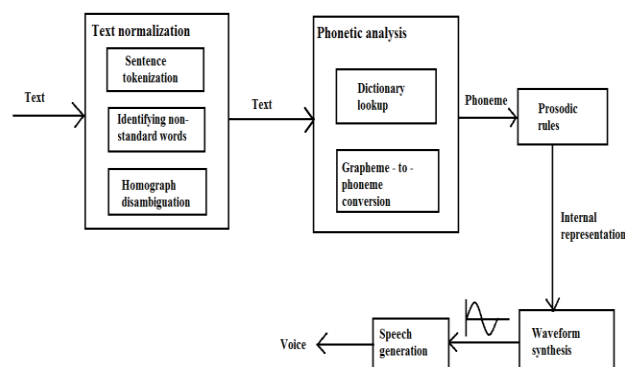


Fig. 2 Block diagram of text to Speech (TTS) model

The above block diagram depicts the outline of how recognition system works. The system consists of detection component, feature extraction components, speech recognition components, Acoustic and language model. The voice recorder reads the analog voice signal supplied by the input and converts it to basic binary code by the ADC (Analog to digital converter) built into the system. The sound recording and word detection components is responsible for taking input from microphone and identifying the presence of word. Word detection is done using energy and zero crossing rate of the signal.

#### IV. FUTURE SCOPE

We are improving the system so as to make the visually impaired people work profusely on their own. Some areas such as web pages translation and verbosity needs to be taken care of. Also, problems like, screen readers reading out unnecessary links, punctuation marks etc. this increases the complexity of the users to understand is to be eliminated without affecting the working of the applications. We need to integrate high speed keeping in mind the cost of the system. Also, the image translation requires more efficient algorithms in order to make a proper description of the images so that people who cannot see can understand easily.

#### V. CONCLUSION

The system we have proposed is not only useful to blind but also to handicapped people. With proper amount of time and resources we will be able to offer visually impaired people low cost, high speed and proper functionality computer system that will benefit the day to day lives of the people and will make them self-reliant.

## ACKNOWLEDGMENT

The success of any paper is not limited to the group of individuals undertaking the paper, it is collective effort of the people around the individual that spell success. There are some key personalities involved, whose role has been very special and vital to pave way to success of the paper. We take the opportunity to express our sincere thanks and gratitude. We are indebted to Dr. M. A. Pradhan, our guide who was very generous in providing us the support we needed. Her invaluable suggestions and time has helped us to make our work possible.

## REFERENCES

- [1] WebAIM(2014, January). "Screen Reader User Survey" [ONLINE]. Available: <http://webaim.org/projects/screenreadersurvey5/>
- [2] C.Asakawa and B.Leporini, "Screenreaders" Chapter of the book UniversalAccessHand book, Stephanidis C. (Ed), Taylor & Francis, June 11, 2009
- [3] Prof.Reeta Singh "Blind Handicapped Computers?" International Journal Scientific and Engineering Research Volume 3, Issue , April-2012 ISSN 2229-5518
- [4] Screenreader-Wikipedia[Online]. Available: <https://en.wikipedia.org/wiki/Screenreader>
- [5] Jagtap Nilesh, Pavan Alai, Chavhan Swapnil, Bendre M.R. "Voice Based System in Desktop and Mobile Devices for Blind People". International Journal of Emerging Technology andAdvanced Engineering Certified Journal, Volume 4, Issue 2,February 2014 ISSN 2250-2459
- [6] JAWS Headquater[Online]. Available: <http://www.freedomscientific.com/JAWSHQ/JAWSHeadquarters01>
- [7] Saravanan, M., and G. Vineet. "SWAPN: A dream tool for the visually challenged." In Global Humanitarian Technology Conference-South Asia Satellite (GHTC-SAS), 2014 IEEE, pp. 87-92. IEEE, 2014.
- [8] Sharma, Kamlesh, and Dr TV Prasad. "Swar The Voice Operated PC." arXiv preprint arXiv:1401.3519 (2014).
- [9] Dobrišek, Simon, Jerneja Gros, France Mihelič, and Nikola Pavešić. "HOMER: a voice-driven text-to-speech system for the blind." In Industrial Electronics, 1999. ISIE'99. Proceedings of the IEEE International Symposium on, vol. 1, pp. 205-208. IEEE, 1999.
- [10] Sharma, F. Reena, and S. Geetanjali Wasson. "Speech recognition and synthesis tool: assistive technology for physically disabled persons."International Journal of Computer Science and Telecommunications 3, no. 4 (2012).
- [11] Ran, Lisa, Sumi Helal, and Steve Moore. "Drishti: an integrated indoor/outdoor blind navigation system and service." In Pervasive Computing and Communications, 2004. PerCom 2004. Proceedings of the Second IEEE Annual Conference on, pp. 23-30. IEEE, 2004.
- [12] Nilesh, Jagtap, Pavan Alai, Chavhan Swapnil, and Bendre MR. "Voice Based System in Desktop and Mobile Devices for Blind People."International Journal of Emerging Technology and Advanced Engineering (IJETA) (2014): 404-407.