

Digital Assistant For The Blind

Prince Bose^{*}, Apurva Malpathak[†], Utkarsh Bansal[‡], Ashish Harsola[§]

^{*}Fr. Conceicao Rodrigues Institute of Technology

400703 Vashi

Email: boseprince201195@gmail.com

[†]Fr. Conceicao Rodrigues Institute of Technology

400703 Vashi

Email: apoorvamalpathak22@gmail.com

[‡]Fr. Conceicao Rodrigues Institute of Technology

400703 Vashi

Email: ubansal25@gmail.com

[§]Fr. Conceicao Rodrigues Institute of Technology

400703 Vashi

Email: ashishharsola@gmail.com

Abstract—The blind and the visually impaired have little to no internet presence because of the absence of cheap solutions to get them online which can be both, hardware and software. Existing technology used for enabling the blind or visually impaired to use the internet or any digital form of information is dependent on Braille displays and keyboards which are expensive and scarce. Another shortcoming of existing technology is that out of all the visually impaired population, less than 2% know how to interpret Braille. Hence a voice controlled system for the blind and the visually impaired was designed, which transceives information in the form of audio. It enables the user to receive and send emails, access daily news, weather forecast, set reminders and alarms and make notes.

Keywords: Voice assistant, Speech Recognition, Low cost, Internet, Speech Synthesis, Blind, Visually Challenged, Raspberry Pi

I. INTRODUCTION

Existing technology used for enabling the blind or visually impaired to use the internet or any digital form of information is dependent on Braille displays and keyboards which are expensive and scarce. Another shortcoming of existing technology is that out of all the visually impaired population, less than 2% know how to interpret Braille. M. Ramirez (2016) made an attempt to make an automatic speech recognition (ASR) system which, along with a haptic interface, aimed at helping preschool children to learn Braille. The algorithm was tested on 9 subjects and its sensitivity was measured in terms of the percentage of true positives. The highest accuracy values were obtained for the a, e, o and u vowels (with hit ratios of 88.8% in all cases), whereas the i vowel exhibited the lowest sensitivity (77.7%). Thus, the drawback of this system was the low accuracy for particular vowels. Another drawback was that this project intended to help teach braille to only the schools in Columbia. A. M. D. Celebre (2015) tried implementing a home

automation system through Siri's capability of speech recognition and through Raspberry Pi as a low-cost control system to automate home devices, namely the air cooler, door, lights, TV and window. A total of 34 subjects with different English accents tested a total of 14 voice commands with a total of 390 trials. The tests verified that the system's average latency is at 2.12 seconds with an overall success rate at 93.3333% in at least 3 trials and 87.381% success rate in at least one trial. In this project, the developers used Apple Siri speech recognition software for recognizing the commands given by the user. The drawback in this paper was that Siri is not an open source software. It is owned by Apple Inc. A. Mishra (2015) aimed at developing a voice-controlled personal assistant robot. According to his approach, the human voice commands are given to the robotic assistant remotely, by using a smart mobile phone. Thus, the drawback was, increase in hardware with a smartphone as well as a microcontroller. Another drawback of this project was that the two devices needed to be connected over a Bluetooth connection at all times, leading to power wastage. Hence we designed an integral system that overcomes all the above flaws. Our proposed system can help everyone of all ages and from all sects of the world, use the Google Speech Recognition API which is open source and available for development purposes and provides the user with minimal hardware and maximum functionality. The objective of this project is to design a voice controlled system for the visually impaired, which will transceive information in the form of audio and enable the user to:

1. Receive and send emails
2. Access daily news
3. Weather forecast
4. Maintain a personal diary / Online Blog

A. Literature Survey

- Patrick Roth , Lori Stefano Petrucci , André Assimakopoulos, and Thierry Pun proposed an audio-haptic feedback mechanism for helping the blind visualize what is on the screen.
of the Association of the Slovenian Blind and Visually Impaired Persons.
- Simon Dobrišek(2003) published a survey paper that covers the evolution of architectures of the information-retrieval systems used in building systems which aid the blind and the visually-impaired people.
- Dr. Hersh Johnson's "Assistive Technology for Visually Impaired and Blind People" covers history of assistive technology for the blind along with block diagrams and approaches used for their implementation.

II. METHODOLOGY

This paper discusses the implementation of a system aimed at helping the blind access online facilities. The system uses speech for communicating with the user. Speech recognition is used to convert the speech input to text. This text is then fed to the central processor which determines the nature of the command and calls the relevant script for execution.

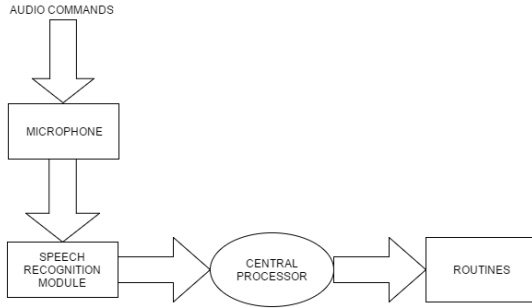


Fig. 1. Basic Block Diagram of our Proposed System.

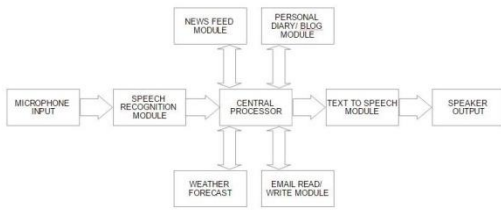


Fig. 2. Complete Block Diagram of our Proposed System.

A. Speech Recognition

The system uses Google's online speech recognition system for converting speech input to text. The speech input

- Jerneja Zganec Gros(1999) proposed HOMER : a voice-driven text-to-speech system developed for blind or visually impaired persons for reading the Slovenian texts. Users can obtain texts from the special corpora organized on the computer network server at the information center from the microphone is temporarily stored in the system which is then sent to Google cloud for speech recognition. The equivalent text is then received and fed to the central processor.

B. News Feed

Upon receiving "National news" or "International news" as the input from the speech recognition module, the central processor calls the script which gathers top news from the Times of India (TOI) RSS feed for national news and Cable News Network (CNN) RSS feed for international news. News reports are then fed to the Text to Speech module which outputs news in form of speech for the user to listen.

C. Email Read / Write

The system periodically downloads and stores the emails the user receives. The user can access his emails using voice commands. The central processor then inputs data from the email read/write module. This email text is then forwarded to the text to speech module to be read aloud by the system. The user can dictate an email to the system. The input speech is converted to text by the speech recognition module which feeds text to the central processor. The central processor then forwards the text to the email read/write module which sends the email with proper formatting.

D. Personal Diary / Online Blog

The user can maintain an offline diary or maintain an online blog. The stored personal notes can be uploaded online on a designed website or stored offline for personal access. The various sections of the note- title, subtitle and body are XML tagged.

E. Weather Forecast

The user can access the weather information about any location for the present day and also the coming 4 days. This information is received in the form of a dictionary which is parsed by the central processor to give user interpretable output.

III. RESULTS

A. Speech to Text Module:

In this module, the speech recognition module was tested. The test was done in a very noisy environment and

the offline method was implemented. The inputs were “Hi! Do I know you?”, “Revenue”, “Hello”, “ABCDEFGH”, “A”, “B”, “C”, “D”, “goodbye”.

Fig. 3. Implementation of Speech to Text Module

B. Text to Speech Module:

The aim of this testing was to check the different voice options we have and to optimize the rate of speech by the D.A.B system for clear hearing and understanding. The inputs were given as text as seen in the code in Fig 3. The output had three voices: male, female and another male. We varied the rate of speech from 70-120 and found 100 as the optimum speech rate.

Fig. 4. Implementation of Text to Speech Module

C. Speech Recognition Module:

In this module, we successfully integrated the TTS and STT system. The given input was speech which after recognition was again spoken back correctly. The online method was implemented for this test.

Fig. 5. Implementation of Speech Recognition Module

D. News Feed Module:

The aim of testing this module was to aptly retrieve the RSS feeds from one of the two source options : Times of India or CNN. The url of times of India was <http://timesofindia.indiatimes.com/rssfeeds/-2128936835.cms> and that of CNN was <http://rss.cnn.com/rss/edition.rss> . A choice was given to the user upon which the user had the choice to opt from either of the two news feeds. There was a small delay in generating the output because of the instability of the internet connection.

Fig. 6. Implementation of News Feed Module - Times

Fig. 7. Implementation of News Feed Module - CNN

E. Google Search Module:

Using this feature we were successful in obtaining all the urls having any correlation with the input word. The input here was “William Stallings”. Thus all links that would be obtained on a regular google search was given to the user.

```
Python Interpreter
*** Python 2.7.10 (default, May 23 2015, 09:40:32) [MSC v.1500 32 bit (Intel)] on win32. ***
*** Remote Python engine is active ***
>>>
*** Remote Interpreter Reinitialized ***
>>>
A moment of silence, please...
Set minimum energy threshold to 57.0180864832
Say something!
Got it! Now to recognize it...
You said search
What would you like to search for?
Got it! Now to recognize it...
Searching for William Stallings
http://williamstallings.com/
http://williamstallings.com/OperatingSystems/
http://williamstallings.com/ComputerOrganization/
http://williamstallings.com/Cryptography/
http://williamstallings.com/DataComm/
http://williamstallings.com/Bibliography.html
https://en.wikipedia.org/wiki/William_Stallings
http://www.amazon.in/William-Stallings/e/B000APXR9Q
```

Fig. 8. Implementation of Google Search Module

IV. CONCLUSION

In this paper, we discussed the design and implementation of a Digital Assistance for the Blind system. The project is built using open source software modules with robust community backing which can accommodate any updates in the near future. The modular nature of this project makes it more flexible and easy to add additional features without disturbing current system functionalities. We have successfully completed building the Speech Recognition module, News Feed module and the Google Search Module. Currently, we are working on porting our system onto our microcontroller Raspberry Pi Model B+.

V. FUTURE SCOPE

- Navigation for the visually challenged using voice instructions and haptic feedbacks can be added for improving the existing system.
- Home automation using microcontrollers can be included in the project.

REFERENCES

- [1] P. Roth, L. Petrucci, T. Pun, and A. Assimacopoulos, “Auditory browser for blind and visually impaired users,” CHI ’99 extended abstracts on Human factors in computer systems - CHI ’99, 1999
- [2] S. Dobrišek, “Analysis and Recognition of Phones in Speech Signal”, 1999, University of Ljubljana.
- [3] Lawrence Rabiner, Bing-Hwang Juang, “Fundamentals of Speech Recognition”
- [4] J.A. Brabyn, “Developments in Electronic Aids for the Blind and visually Impaired”, IEEE Eng. In Medicine and Biology Mag., Dec. 1985. Pp. 33-37
- [5] Elsevier, “Human computer interaction: proceedings of the Fifth International Conference on Human Computer Interaction”, (HCI International ’93), Orlando, Florida, August 8-13, 1993

- [6] Hersh, Dr Marion, “Assistive Technology for Visually Impaired and Blind People”, Springer ISBN: 9781846288661
- [7] Information Science Reference, “Human computer interaction: concepts, methodologies, tools, and applications”, 2009
- [8] McGraw-Hill, “Human computer interaction: psychology, task analysis, and software engineering”, 1992
- [9] Bonn, “Human computer interaction: from individuals to groups in work, leisure, and everyday life ; proceedings”, Germany, September 5-8, 1994
- [10] Computer Assisted Learning, “Computer Assisted Support For Blind Students: The Use Of A Microcomputer Linked Voice Synthesizer”, 1981
- [11] International Conference on Acoustics, Speech, and Signal Processing, “Computer assisted voice identification system (CAVIS)”, ICASSP-88
- [12] M. Bansode, S. Jadhav, and A. Kashyap, “Voice Recognition and Voice Navigation for Blind using GPS,” Ijireeice, vol. 3, no. 4, pp. 91–94, 2015
- [13] T. Vincent, “Computer assisted support for blind students,” Computers & Education, vol. 6, no. 1, pp. 55–60, 1982
- [14] A. Mishra, P. Makula, A. Kumar, K. Karan, and V. K. Mittal, “A voice-controlled personal assistant robot,” 2015 International Conference on Industrial Instrumentation and Control (ICIC), May 2015.
- [15] A. M. D. Celebre, A. Z. D. Dubouzet, I. B. A. Medina, A. N. M. Surposa, and R. C. Gustilo, “Home automation using raspberry pi through Siri enabled mobile devices,” 2015 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Dec. 2015.
- [16] M. Ramlrez, M. Sotaquira, A. De La Cruz, E. Maria, G. Avellaneda, and A. Ochoa, “An automatic speech recognition system for helping visually impaired children to learn Braille,” 2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA), Aug. 2016.