Implementation of E-mail Access for Vision Impaired Person Using Speech Synthesizer

Rashmi Kurkure  
KBT College of Engg  
Nashik, (India)

Niketa Gagrani  
KBT College of Engg  
Nashik, (India)

Sakshi Gupta  
KBT College of Engg  
Nashik, (India)

Sonali Phadol  
KBT College of Engg  
Nashik, (India)

ABSTRACT: Developments in technology over the last twenty years have enhanced the possibilities for learning and working for people with vision impairments. Computers with speech or Braille output peripherals, or portable speech-output PDA devices, or GPS-controlled tactile navigational maps, are some examples.

In this paper we have proposed a system that will help the people with vision impairment to access their mail accounts conveniently. We are using Mail Synchronization and Retrieval, Text to Speech, Speech to Text, Screen Magnifier technique. Through this Application the Vision Impaired user can get a speech synthesized output of the mails in his inbox or can view his email using the magnifier option available.

Keywords: Mail Synchronization and Retrieval, Screen Magnification, Text to speech recognition, Speech to Text.

1. INTRODUCTION

Social media is more than a buzzword. Social media is a broad term used to describe different types of web based services and mobile applications that enable people with common interests to communicate and share information and resources with each other in real time. Web accessibility means that people with all abilities and disabilities can use the web, that means, can perceive, understand, navigate, and interact with the web. When web applications are correctly designed, developed and edited, all users can have equal access to information and functionality, also they can be accommodated without decreasing the usability of the application for non disabled users. The most common and essential need of internet is accessing emails[1].

It is difficult for visually impaired persons for accessing any kind of websites because of their significant limitation of visual capability. To overcome from these drawbacks we have proposed a system which will help these users to access their mail.

Visual impairment is the consequence of a functional loss of vision, rather than the eye disorder itself. The terms "partially sighted", "low vision", "legally blind" and "totally blind" are used by schools, colleges, and other educational institutions to describe students with Visual impairments. They are defined as follows:

- Partially sighted indicates some type of visual problem, with a need of person to receive special education in some cases;
- Low vision generally refers to a severe visual impairment, not necessarily limited to distance vision.
- Legally blind indicates that a person has less than 20/200 vision in the better eye after best correction (contact lenses or glasses), or a field of vision of less than 20 degrees in the better eye.

Block diagram of Proposed System:
Through this Web Application the Vision Impaired user can get a speech synthesized output of the mails in his inbox or can view his email using the magnifier option available. To compose emails he can give a voiced input to the speech recognition module which converts the speech dictated by the user to text. Adequate measures have been taken which makes this Application more effective.

2. RELATED WORK DONE

I] Text to Speech Synthesizer

The Text to Speech (TTS) Synthesizer first converts the input text into its corresponding linguistic or phonetic representations and then produces the sounds corresponding to those representations. Speech synthesis is the artificial production of human speech. A computer system used for this purpose is called a speech synthesizer, and can be implemented in software or hardware products.

II] Automatic Speech Recognition

Automatic speech recognition (ASR) can be defined as the independent, computer driven transcription of spoken language into readable text in real time (Stuckless, 1994). ASR is technology that allows a computer to identify the words that a person speaks into a microphone or telephone and convert it to written text. The ultimate goal of ASR research is to allow a computer to recognize in real time, with 100 percent accuracy, all words that are intelligibly spoken by any person, independent of vocabulary size, noise, speaker characteristics or accent.

3. TECHNOLOGY USED

3.1 Microsoft Speech Object Library in C#.net

Using the .Net Framework SDK or the “Add Reference” option in visual studio.NET, you can import the Microsoft Speech Object Library (sapi.dll) into our project. Microsoft Speech SDK enables a developer to add Speech capability into an application. Microsoft .NET framework provides System. Speech. Synthesis for voice synthesis.[2]

3.2 Screen Magnifier

A screen magnifier is software that interfaces with a computer’s graphical output to present enlarged screen content. It is a type of assistive technology suitable for visually impaired people with some functional vision. This application can magnify a portion of the screen up to 16 times. Once the software is activated, the user can go to a specific area in the screen. Immediately after setting the focus, the magnifier will enlarge the text or graphics in that area.

4. PROPOSED WORK

We have proposed system and are implementing an application for vision impaired persons that will help them to access their email accounts. Visually impaired persons today, have to acquire specific skills in order to use the browser effectively when browsing the Internet. Yet, the visually impaired learners are deprived of this very important learning tool.

The visually impaired learners are also deprived of enjoying services in the Internet like sending and receiving e-mails unlike their other normal counterparts. Due to the fact that the conventional browser available is developed for normal users that enable them to control the functionalities designed in the browser, it is thus not suitable for the visually impaired learners.

5. IMPLEMENTATION DETAILS

5.1 Speech Synthesizer

A speech synthesizer takes text as input and produces an audio stream as output. Speech synthesis is also referred to as text-to-speech (TTS). A synthesizer must perform substantial analysis and processing to accurately convert a string of characters into an audio stream that sounds just as the words would be spoken.[4]

Fig 5.1: phases of speech synthesis
5.1.1 Text Analysis

The front end specializes in the analysis of text using natural language rules. It analyzes a string of characters to determine where the words are (which is easy to do in English, but not as easy in languages such as Chinese and Japanese). This front end also figures out grammatical details like functions and parts of speech. For instance, which words are proper nouns, numbers, and so forth; where sentences begin and end; whether a phrase is a question or a statement; and whether a statement is past, present, or future tense.[3]

5.1.2 Phonetic Analysis

It tries to split text into phonemes. It converts grapheme to phoneme conversion (letter to sound).[4]

5.1.3 Prosodic Analysis

It adds prosodic controls like melody, accent, and pauses to the phoneme string.[4]

5.1.4. Sound Generation

The back end has quite a different task. It takes the analysis done by the front end and, through some non-trivial analysis of its own, generates the appropriate sounds for the input text. Older synthesizers (and today’s synthesizers with the smallest footprints) generate the individual sounds algorithmically, resulting in a very robotic sound. Modern synthesizers, such as the one used by the Microsoft Speech Platform Runtime 11, use a database of sound segments built from hours and hours of recorded speech. The effectiveness of the back end depends on how good it is at selecting the appropriate sound segments for any given input and smoothly splicing them together.

We are working on Speech-to-Text to compose mails and screen magnifier for enlarging text and images.[3]

6. RESULT

We have used POP3 at back end for mail synchronization and retrieval. With this we can receive emails and can synchronize servers with the Application.

IMAP and POP are two different protocols. There are many differences between these two. The main difference is that IMAP (Internet Message Access Protocol) always syncs with mail server so that any changes you make in your mail client will instantly appear on your webmail inbox. On the other hand, in POP (Post Office Protocol), your mail client account and mail server are not synced. It means whatever changes you make to your email account in the mail client will not be transferred to the webmail inbox. In simple terms, if you are using IMAP and mark a mail as read, it gets marked as read in your web based inbox too (because the changes are happening on the server). However, this won’t be the case if you are using POP, because the mails are downloaded to your PC and the changes won’t reflect on the server.

So we are switching from POP3 to IMAP. But there are limitations which have to be taken into account. The most restrictive factor is, it consumes lots of time for retrieving all mails from mail server. So we will retrieve only few recent mails from the server.

We have implemented this application using C sharp language, for front end designing, which is intended to be a simple, modern, general-purpose, object-oriented programming language.
Fig 6.2: text to speech synthesizer

After retrieving this text when user click on speak button the related sound is generated. We have tested this application by writing different text for ex: 1,2,3 gives output as one two three and for 123 it gives output as one hundred and twenty three, for Mr. it speaks as mister.

7. CONCLUSION

This application will help the vision impaired user to access independently their email account. This web application not only makes the email access more complete but also could be a harbinger to the new world of Assistive Web Technologies.

8. FUTURE SCOPE

There are several features that can accommodates in implemented web application such as attach files, download files and chat window like features for vision impaired persons. Further we can extend the output of text speech synthesizer by storing generated sound for later use.

9. Acknowledgement

We take this opportunity to thank Prof. Nitin S. Ujgare and Prof. Vaishali S. Pawar for their valuable guidance and for providing all the necessary support to accomplish this research. We would like to extend our gratitude towards our beloved Principal Dr. Jayant T. Pattiwar for his great support.

10. REFERENCES


