Voice Based Email for Blind

Amritha Suresh*, Binny Paulose, Reshma Jagan, Joby George

Department of Computer Science, MG University, Kottayam, Kerala, India

ABSTRACT

E-mail remains the most go-to form of the many communication technologies in the business world. However all these technologies can be of no use to the people who are visually impaired as all activities that can be performed on the computer is based on visual perception. As nearly 285 million people worldwide are estimated to be visually impaired it becomes necessary to render internet facilities for communication useful for them also. This paper aims at developing an email system which helps them to use it without difficulty. The system makes use of speech recognition, Interactive Voice Response (IVR) and mouse click events. As an additional security aspect, the system makes use of voice recognition for user verification.

Keywords : Speech-to-text conversion, text-to-speech, IVR, MFCC, Dynamic time warping, voice recognition.

I. INTRODUCTION

Worldwide email use continues to grow at a healthy pace. In 2015, the number of worldwide email users was nearly 2.6 billion. By the end of 2019, the number of worldwide email users will increase to over 2.9 billion. Over one-third of the worldwide population will be using email by year-end 2019. One of the major drawbacks that sets in is that accessing emails or on a whole any page on the internet requires a person to have visual capabilities. This means that a visually challenged person can in no way take the benefits of the facilities provided by the internet thus rendering the technology useless.

The existing technologies like screen readers, automatic speech recognizers, speech-to-text and text-to-speech which came into picture, made it easier for the visually impaired, though only partially. Voice based email system with these technologies alone pose privacy and security problems for a visually impaired person. So, there emerged a need to create a complete voice based application with which emails can be sent or received providing security and privacy.

Therefore we came up with this system Voice based email for blind which aids visually impaired people to use email facilities with ease. The most crucial aspect we are considering while developing this system is providing security for the user by incorporating speaker verification during sign-in.

The users of this system need not have any basic information about the keyboard shortcuts used or where the keys are located. All functions used in our system are based on simple mouse click operations making the system very user friendly.

II. METHODS AND MATERIAL

1. Existing System

The mail services that are available today are of no use to the people who are visually impaired. This is because these systems are not helpful to them in anyway as it cannot provide any audio feedback to readout the contents for them. As they are unable to visualize things that are present on the screen, they find it difficult to perform operations such as performing mouse click specifically.

Although, there are screen readers available but, they impose some or the other kind of difficulty to them. Screen readers basically read out the content on the

screen for them and in order to respond to it, they need to provide input through a keyboard. So, in order to accomplish this, the user needs to be aware of the positions of the keys on the keyboard. Hence, a person who has never made use of a computer will never be able to use such kind of a system.

2. Proposed System

The most crucial aspect that we are considering while developing this system is that privacy of a user will not be compromised as we incorporate voice recognition for user verification. The other important aspect is that the users of this system need not have any basic information about the keyboard shortcuts used or where the keys are located. All functions used in our system are based on simple mouse click operations making the system very user friendly. The system will also be continuously prompting the user of which click will perform which operation thus making it easy for the user as he/she won't require to remember the operations.

The proposed system incorporates voice recognition to verify the user for logging into the email system. Speaker recognition is the process of automatically recognizing who is speaking on the basis of individual information included in speech waves. It has two sessions. The first one is referred to the enrolment session or training phase while the second one is referred to as the operation session or testing phase. In the training phase, each registered speaker has to provide samples of their speech so that the system can build or train a reference model for that speaker. In case of speaker verification systems, in addition, a speaker-specific threshold is also computed from the training samples. During the testing (operational) phase, the input speech is matched with stored reference model(s) and recognition decision is made.

A voice analysis is done after taking an input through microphone from a user. The design of the system involves manipulation of the input audio signal using MFCC feature extraction. At different levels, different operations are performed on the input signal such as Pre-emphasis, Framing, Windowing, Mel Cepstrum analysis and Recognition (Matching) of the spoken word. The voice algorithms consist of two distinguished phases. The first one is training sessions, while, the second one is referred to as operation session or testing phase.

DTW algorithm is based on Dynamic Programming techniques. This algorithm is for measuring similarity between two time series which may vary in time or speed ie. for feature matching.

This technique also used to find the optimal alignment between two times series if one time series may be "warped" non-linearly by stretching or shrinking it along its time axis. This warping between two time series can then be used to find corresponding regions between the two time series or to determine the similarity between the two time series.

Registration is the first module of the system. Any user who wishes to use the system should first register to obtain username and password. This module will collect complete information of the user by prompting the user as to what details needs to be entered. The voice samples of the user are collected for voice recognition.

Once the registration is done the user can login to the system. This module will ask the user to provide the username and password. This will be accepted in speech. Another voice sample is asked for performing the speaker verification process. The user is redirected to the inbox page once log in done successfully. If the user is not the authorized one, page redirects back to the login page.

Once logged in, the user can perform normal operations of a mailing system. The options available are: Compose, Inbox, Sent mail and Trash. Voice commands are used to navigate between these pages. In compose, user will be prompted to speak the message and send it. In inbox, the user can listen to mails he/she wants to by performing the click operation specified by the prompt. In order to access the sent mails user will need to perform the actions provided by the prompt to navigate between mails. When the control lands on particular mail user will be prompted as who the receiver was and what is the subject of the mail. Deleted mails in trash could be the ones from inbox or sent mail. If at any time the user needs to retrieve a mail which was deleted it can be done from this option.

III. RESULTS AND DISCUSSION

IMPLEMENTATION

Since the system is for the visually challenged people who are naïve to computers or have very little experience in using technologies available to assist them in using the computers we have kept continuous prompting done by the system in language understandable to them. The user will just need to follow the instruction specified and perform actions according to their need. As the users are considered naïve to computer systems, we are not keeping any functionality in which use of keyboard related operations are used. The complete website will work on just mouse click operations and speech commands.

A. Registration

Registration is the first module of the system. Any user who wishes to use the system should first register to obtain username and password. This module will collect complete information of the user by prompting the user as to what details needs to be entered. Once the registration is done the user can login to the system. The user will need to speak up the details to which the system will again confirm by prompting alphabetically and these details are stored in database for speaker identification. If the information is not correct user can re-enter else the prompt will specify the operation to be performed to confirm.
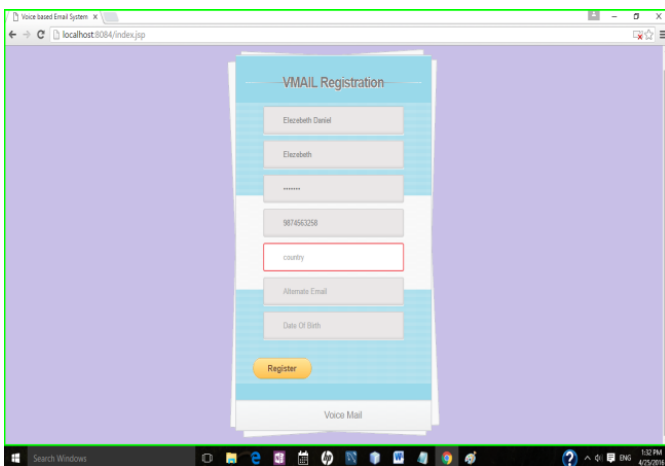

B. Login

This module will perform authentication check every time the user would require accessing his/her account. It will accept username and password in speech format and convert it to text. This text will then be used to authenticate the users. Once the user is found authentic s/he is redirected to their respective homepage. This authentication is done using speaker identification by feature extraction, identification and matching. Only a registered user can login to the system.
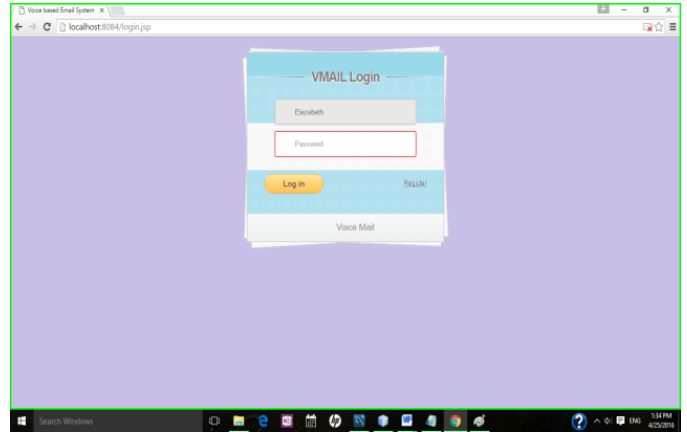

C. Compose

This is the most important module of the complete system. Once in this module user can create the mail he/she wants to send. The main difference between the existing system and our system is that unlike other normal system our system will ask for recording and the recorded audio file will be sent to the other end as an attachment. The user will be prompted continuously about the mouse click operations that need to be performed to access the system.
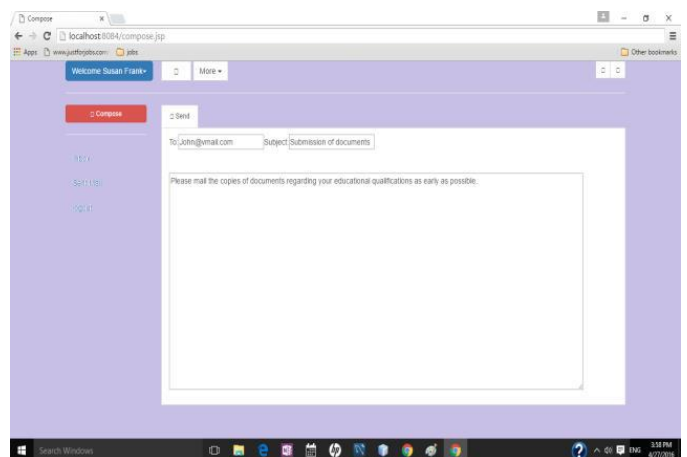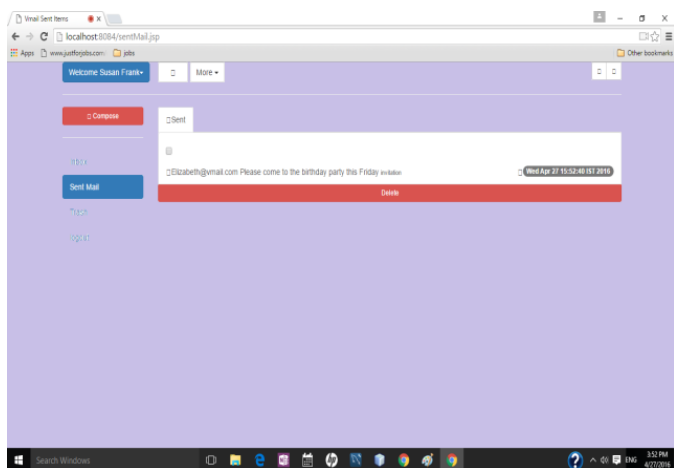
E. Inbox

The user can view all the mails received to the account through this option. Once in this option the system will be continuously prompting the user about which click operation needs to be performed to navigate through the page and perform operations on the received mail. The user also has the option to delete the mails received. The deleted mails will be stored in the trash section.
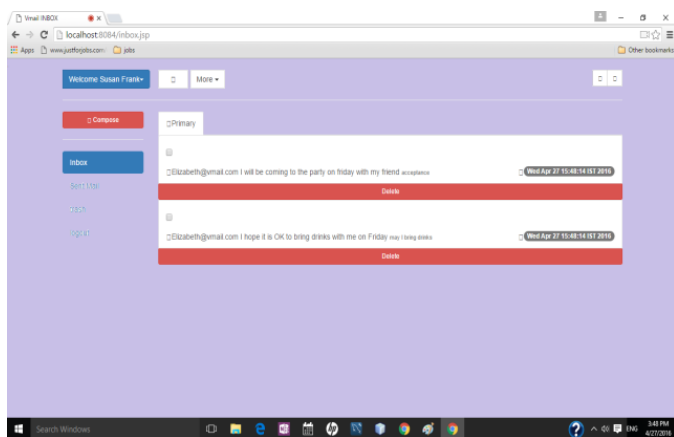

F. Sent Mail

The mails sent by the user will be stored in this option. This will help the user access mails that were sent by him/her. User will be prompted as to which mouse operations are to be performed in order to gain access to specific operations. As the user navigates between mails he/she will be prompted about the receiver and subject of that mail. This will provide ease of access to users.

G. Trash

The mails that will be deleted by the user will be saved in this section. User can delete mails from both inbox as

well as sent mail option. This option will help the user to retrieve mails that were deleted by user at that time but are required currently.
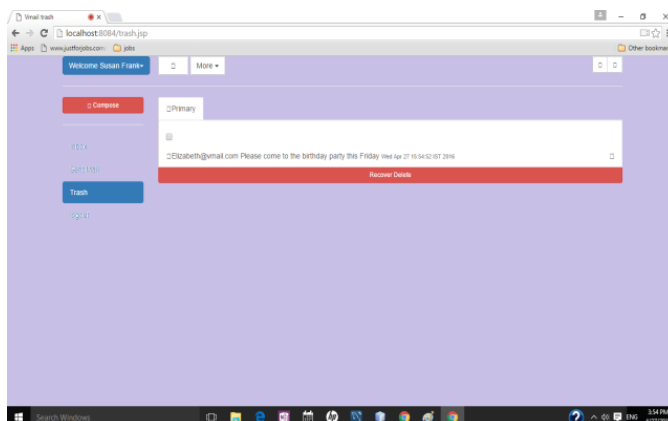
IV. CONCLUSION

The main reason for developing this system described in the paper is to create ease for visually impaired people to use the most pervasive form of communication in today's world; email. This mail system will help overcoming all the minor difficulties which are faced by visually impaired people as the system works on vocal abilities. This will reduce the software load of using screen readers and automatic speech recognizer and also the user's cognitive load of remembering keyboard shortcuts. The system will be guiding the user as to what operation needs to be performed for obtaining desired results through IVR thus making the system much user friendly. The system developed now will be working only on desktops. As use of mobile phones is emerging as a trend today there is a scope to include this facility as an application in mobile phones also. Also security features to be implemented during login phase can be revised to make the system more secure.

V. REFERENCES

[1] Jagtap 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 (IJETAE), 2014 on Pages 404-407 (Volume 4, issue 2).

[2] Ummuhanysifa U.,Nizar Banu P K , "Voice Based Search Engine and Web page Reader". In

Internationa Journal of Computational Engineering Research (IJCER). Pages 1-5.

[3] 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).

[4] Shabana T., Anam A. , Rafiya A. , Aisha K. "Voice Based Email System for Blinds". International Journal of Advance Foundation And Research In Science & Engineering (IJAFRSE) Volume 1, Issue 10 , March 2015.

[5] Priyanka Mishra, Suyash Agrawal. "Recognotion of Speaker Useing Mel Frequency Cepstral Coefficient & Vector Quantization for Authentication". International Journal of Scientific & Engineering Research Volume 3, Issue 8, August-2012.