



VOICE BASED SYSTEM

Riya Jose¹, Rosemol Joseph², Pinky Jose³
Livya George⁴

⁴Assistant Professor, ^{1,2,3,4}Computer Science, Sahrdaya College of Engineering and Technology

Abstract

The paper deals with creating a human computer user interface that has speech recognition technology. It involves modifying the user experience of applications such as the file manager, browser, word processor and web applications such as email and other social sites such as facebook, twitter etc. It helps to provide maximum action through speech and minimal user input through other hardware. So, our project will be very helpful for the paralysed persons who have an intense desire to use the computer system for their needs. It aids them to reach their favorite folders and make updates to the existing folder through speech. Through our project, the entire folder content can be read out loudly on users request so that they can hear and understand the content of the particular folder.

INDEX TERM: Speech recognition, Speech synthesis, File manager, Speech interfaces

I. INTRODUCTION

A computer operated by user's voice implements the commands given by the user as computer actions. User data are stored in files so that it can be used whenever needed by corresponding applications. The files may contain audio, video or simply a text document. The focus of building this paper is to implement a system that makes the ultimate advantage of speech control using computer operating method. It is really convenient to operate a system by our own voice. One of the main challenges faced by the system is false voice recognition. Through speech synthesis the system outputs the results using user interfaces of natural language commands.

Apart from file management, this paper also focuses on opening browser on user commands, setting alarms, whether forecasting. System sets alarm on user command. On being instructed by the user, browser will be opened for necessary functionalities. System outputs the results through speech synthesis by using natural language command as interface. Each and every file is very much vital for the user so is selection and manipulation of it. The interface between user and file is done by the file manager.

The users of this system need not have any basic knowledge regarding any keyboard shortcuts or where the keys are located. All functions are based on voice based operations making it very easy for any type of user to use this system. Also the user need not worry about remembering which operation he/she needs to perform in order to avail system service.

II. RELATED WORKS

There were many efforts in the direction for creating a speech recognition system aiding file retrieval. Vishnudas Raveendran, Mary Pauly, Nicky Paul [1] proposed a system which enables database interaction using speech input. This system uses MySQL as the backend to store file attributes and details based on the extracted information and the action required, an SQL query is built dynamically and will be executed. There will be tag-id and file-id fields in the database. The tag-id obtained will be checked with the file-id to obtain each file associated with the tag-id. If there are more than one tag-id that is matching for a file then it will be ranked higher in search results. The file-id will be then

used to retrieve the file attribute and hence the file.

S.Nareshkumar,N.Mariappan,K.Thirumoothy [2] proposed a system using sphinx4 to take a voice input from the user and then converting it to text.sphinx4.Text.sphinx4 is a HMM based speech recognizer.These voice inputs will be stored as wave file which will be put together with the configuration file and will be later processed to get a text file.The text obtained will be then converted into a standard text format.Every database has a corresponding configuration XML file holding all details about itself and its content.The text is initially parsed to return tokens.The stop words are then removed from these tokens and categorized under different labels.There is table containing the patterns and corresponding conversions to be used for mapping the text to SQL query.The query thus obtained is used to interact with thw database inorder to retrieve requested data.

III. DESIGN PRINCIPLES

A. Need for the system

the same name in another directory.Thus there is a element of disorganisation in the hierarchical organisation of files.

B. System Architecture

The overview of the system architecture is shown in figure 1.The flow of the system starts when the user gives a command using speech recognition(SR).After recieving the transcript from the SR, it will be passed to the custom language processor.It will split the command into individual words and text preprocessing will be done to remove the stop words and lemmatize the remaining words.The keywords will be used for executing the corresponding actions.Once the required details of the command is found, the system will be executing the actions.The final results will be spoken to the user.

Trouble physically using the keyboard - Voice input systems can allow a person to operate a computer without using a keyboard or mouse. This can help people who are not able to use their hands at all, as well as others who can use their hands but are limited by speed, fatigue, or pain (e.g., carpal tunnel syndrome, one-handed keyboarders).

Trouble creating text - Voice input systems can help a person, who has difficulty spelling words, create text. This can be particularly useful for some people with learning disabilities. However, the person's reading abilities need to be strong enough to recognize when the computer displays the wrong word.

Existing file managers are based on hierarchical structure of directories.There exist a need for the users to remember the path to the files.Apart from the visual elements to display the files and folders in each directory the user is traversing.Also a file or a folder may have a unique name in a particular directory however there also exist a chance that the user meets

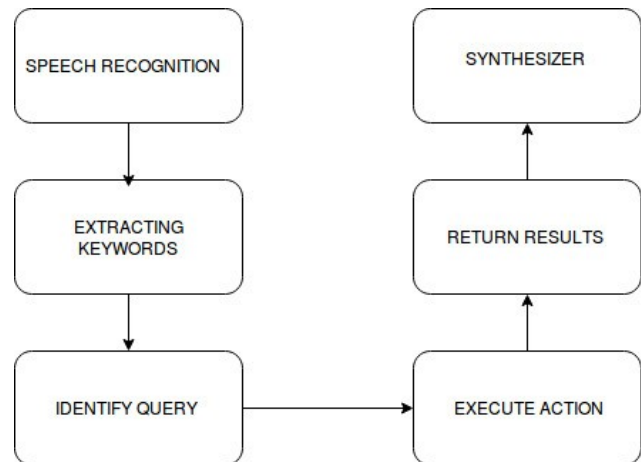


Fig. 1. Flow chart of the system

IV. MODULE DESCRIPTION

The modules that are required for implementing voice based system are

A. Speech Recognition

Speech recognition 1.2.3 it is the library for performing speech recognition with the google speech recognition API. The Secure Base Operating System (SBOS) recognizes each command with the help of this module. Speech recognition API records the audio spoken by the user and encodes it using Free Lossless Audio Codec (FLAC) encoding after which data is sent to the google speech recognition.

B. Command Generation

The user application allow the user to perform tasks and use the computer for various tasks and for various applications like accessing E-mail using an E-mail client, listening to music using browsing the internet . all these applications in our system have been tailored to be used for speech interfaces

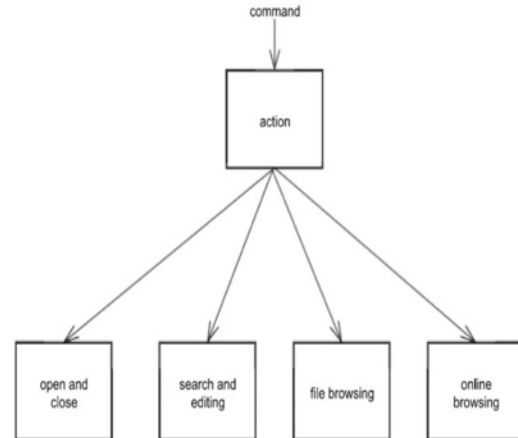
. Hence the user can execute the tasks using voice commands

.In our approach the user is freed from the responsibility of travelling through the folders and provides a method to retrieve files based on the various file attributes and relevance . the project focus on the design and architecture of the file manager . the advantage of using this system is that it is specifically for speech interfaces and allows the user to use the system with minimal or no displays . In computing systems, an file is the stored data of users in a form which can be accessed by its respective application . it can be a document , image, audio, video or any application specific document

C. Action

The command is recognized and identified and it is pro- ceeded for further action. If the command was to open a file , first it will check if it is from a authenticated user then the instruction is analyzed and understood . Action is performed in the final stage ie, opening the file requested. It is high time to introduce a file manager which is designed to work with speech interfaces . however , speech interface allow to directly invoke a file's name instead of travelling through folders. Through this project , we approach to designing file

manager for speech interfaces. The user interface with the system through natural language commands and the system outputs the



results through speech synthesis.

Fig. 2. Output action

v. RESULTS AND DISCUSSION

The system was surveyed among the users proficient in using the computers. The survey involved explains the basic concepts and functionalities of the system along with some demonstration of simple tasks. Users where then encouraged to try using thier own commands and then performing a list of sample tasks. The observations are given below

1) Most of the commands which were interpreted by the system was correct with very low missrate.

2) Some users find the need to use GUI or some text representation of the spoken command.

3) A change from the GUI interface to speech interface faced some initial misscomfort even though the users felt comfortable once got used with the system.

4) The users gave an encoringing response for the hands free approach of the system.

V. CONCLUSION

The system is in the initial prototype was found acceptable among the users for managing the computer system through voice. Users felt at ease in using an interface as natural as speech recognition and synthesis. The system has its short coming in the form of providing a few visual representation to the auditory system. Th e visual representation

would have complimented the users miss of attention for listening the speech synthesizer. However it is proved to be a better computer navigation system for speech interfaces.

VI. FUTURE SCOPE

The clustering, context analysis and summarization can be used in machine learning tools that can benefit the system. In order to support environments with poor internet connectivity offline speech recognition tools such as pocketsphinx can be used.

VII. REFERENCES

- [1] An Approach To File Manager Design For Speech Interfaces Vishnudas Raveendran, Mary Pauly, Nicky Paul, Priya S, Raisa Varghese Assistant Professor Computer Science Dept. Sahrdaya College of Engineering and Technology Thrissur, India
- [2] N.K.S. Nareshkumar, "Database interaction using automatic speech recognition", in International conference on innovations in engineering and technology, 2014
- [3] M.L.R.M.T. Bayoumi "SPHINX: A high level synthesis system for DSP Design" in IEEE International Symposium on Circuits and Systems, 1992
- [4] Microsoft, "Common speech recognition commands", Microsoft, [Online]. Available: <https://windows.microsoft.com/en-in/windows/common-speech-recognition-commands#ITC=windows-7>.
- [5] P.s foundation, "Python", Python software foundation [online]. Available: <https://www.python.org/>.