

An Interactive Computer Environment Using Speech Recognition

^{#1}Purva shelar, ^{#2}Sushrut Singatkar, ^{#3}Megha Tangade, ^{#4}Vijay Varma

¹purvas3@gmail.com

²mail.sushrut@gmail.com

³meghatangade567@gmail.com

⁴varma.vijay69@gmail.com



^{#1234}Department Of Computer Engineering,

TSSM's P.V.P.I.T Collage Of Engineering Pune, India

ABSTRACT

This paper presents an interactive interface that uses speech as a new input modality to provide user-friendly access to computer. Controlling applications is a main part of automation, the main objective of this work is to make a computer system which is controlled by voice commands by user. This system provides a way that the blind and physically disabled population can easily control many functions of a computer via speech, this system also provides blind people with a function via TTS (Text To Speech) if they want to receive contents of an e-mail. The paper focuses on implementation details and structure of the whole system to make it extensible and adaptable for various applications to be added to this system.

Keywords: Control by speech, Human Computer Interfacing, TTS, STT, SAPI interface, Robot API.

ARTICLE INFO

Article History

Received :25th October 2015

Received in revised form :

27th October 2015

Accepted:28th October , 2015

Published online :

2nd November 2015

I. INTRODUCTION

With the rise of the technical inventions computer has become part of our lives, however, using the typical input method with keyboard and mouse has a great possibility bringing various syndromes that influences health. This system is used to provide convenience for user to control and monitor the Device. Imagine a system which downloads an important e-mail and also read it for you.

Therefore a more flexible and healthy operating method should be developed instead of the traditional one. There are certain solutions that can improve traditional computer operation method:

- Voice-Control Computer, which based on a human computer interface able to convert voice commands into computer actions.
- Using voice command perform basic PC (Personal Computers) surfing operations.
- Voice-Email client, using voice command download, read and preview E-mail.
- Voice-Media Control, using voice command control basic media control operations.

As by the growth of PC (personal computers), technology makes it easy for a user to access and controls the device by speech. This system defines the use of speech reorganization in order to control the operations going on the desktops. It is mainly focused on the handicapped people, and for assist dyslexic students who have problems with character or word use and manipulation in a textual form they can simply give a voice command to handle the operations performed on system using voice.

Taking into consideration the growing demand for speech enabled system this proposed work will help tremendously to access web. The user will say a command through the microphone, this command will be converted into text with the help of Microsoft's SAPI (Speech Application Programming Interface), then command will be mapped with database information and the appropriate actions will be taken.

II. RELATED WORK

A) SAPI

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 the Speech API is a freely redistributable component which can be shipped with any Windows application that wishes to use speech technology. Many versions (although not all) of the speech recognition and synthesis engines are also freely redistributable.

There have been two main 'families' of the Microsoft Speech API. SAPI versions 1 through 4 are all similar to each other, with extra features in each newer version. SAPI 5 however was a completely new interface, released in 2000. Since then several sub-versions of this API have been released.

B) Robot API

The ability to control mouse pointer position, motion and clicks is not available in MATLAB. To work around this issue, you can use the Java class `java.awt.Robot` which has this ability. The Java documentation describes the purpose of the `Robot` class as follows: This class is used to generate native system input events for the purposes of test automation, self-running demos, and other applications where control of the mouse and keyboard is needed. This class has three main functionalities: mouse control, keyboard control, and screen capture. Here are some of the important member functions:

- **Mouse control functions**

`void mouseMove(int x, int y)`

This function moves the cursor to the coordinate (x, y) which is defined with respect to the top-left screen corner (in contrast to Matlab's coordinate origin at the bottom-left corner).

`void mouseRelease(int buttons)`

This pair of functions performs the button click. Their input argument is an OR'ed combination of `java.awt.event.InputEvents`.

- **Keyboard control functions**

Keyboard action is emulated by the following pair of functions. Their key codes are defined in `java.awt.event.KeyEvent`:

`void keyPress(int keycode)`

`void keyRelease(int keycode)`

III. EXISTING SYSTEM

This paper handles remote control computer system using speech recognition technologies of mobile devices for the blind and physically disabled population. The configuration of the system consists of a mobile device such as a smartphone, a PC server, and a Google server that are connected to each other. Users can command a mobile device to do something via speech such as directly controlling computers, writing emails and documents, calculating numbers, checking the weather forecast, and managing a schedule. [2] In this proposed work the main focus is on the WEB APPLICATIONS. This system will help them to download news, or even access their mails through speech. The idea is to incorporate several applications like Email Reader/Sender, News, Reader, Web Content, Blog, RSS Reader, Local System File Reader, Text/Document Reader, and Voice Command System. This is an attempt to develop web application through speech interaction. [3] In this paper, the design and evolution of a mobile sound transcription tool for the deaf and hard-of-hearing is elaborated. Transcriptions include dialog and descriptions of environmental sounds. The transcriber is a multilingual transcriber who listens to the audio and sends the converted text message to the server, the server then using the IP address of the user sends it to the user. If the user is not logged into the server then the message is stored in the server database and sent when the user logs in. Authentication of user is done via username and mobile phone number, whereas for transcriber it is username and password to have a secure environment.

IV. PROPOSED SYSTEM

The main goal of the system are-

A) Voice Controlled Mouse pointer

Controlling mouse with human voice is main motive which will provide the benefit to the people with disabilities to use computer in more effective and efficient way.

B) To Use Speech And Non-Speech Characteristics Of Vocal System

Existing systems are developed using speech or on the word recognition, proposed system will provide extra functionality to the system by controlling mouse pointer using vowel sound and cursor's direction and velocity will be decided by the vocal characteristics such as pitch, energy, loudness etc. Using both speech and vowel sound commands to the system, needed action should be performed. Our main concern will be with non-speech characteristics of human vocal system to perform operating system operations. Because limited research has been done in the area of non speech characteristics of human voice so we will use non speech voice characteristics such as pitch, vowel sound and energy to efficiently commanding computer to perform operations for example:

1. Double click.
2. Dragging.
3. Picking file from one folder and drop to another.

4. Continuous path following (e.g., in games and drawing).
5. Scrolling.

C) Voice Controlled Multimedia

Controlling Multimedia with human voice is feature which will provide the benefit to the people so that they can control Multimedia functions like Volume up , Volume down , pause track , next track and previous track.

D) Controlling Email Client via Mail

This project provides an innovative feature via which we can download emails and read them out using STT (Speech-To-Text). STT is a system which recognizes and gives the output in the form of speech.

V. SYSTEM ARCHITECTURE

In the proposed architecture the speech input is given through the microphone to the computer. Microphone processes the audio stream to the Speech Recognition system which will convert a speech signal to a sequence of words in form of digital data which can also be called as a command. As mentioned above the SAPI interface will mainly use for conversion of this speech to text.

Command comparison search provides a better match to the user’s current needs. Action mapping process involves mapping the keywords retrieved from context search with a particular action. This process will result into completing the appropriate action exactly what the user wanted. This action can be known as an event and can be any of the E-mail client, mouse control, PC surfing, media control Interface. After that the result is sent to SAPI which will convert this text to again speech so that the user would be able to hear it.

For accuracy purpose the command based search is used for keywords that user enters so that appropriate action can be taken when appropriate keywords are encountered. User must be able to set/change the system preferences and context search parameters as per the needs. Even multilingual commands would be accepted and recognized. For this the system must allow the user to map standard actions/commands to new keywords (keywords from different languages). System must also allow the user to create new actions and map respective voice commands to them. Using advanced Robot API, the system would generate mouse, keyboard events so that almost any application can be controlled using this system.

Following shows the overall flow of the system.

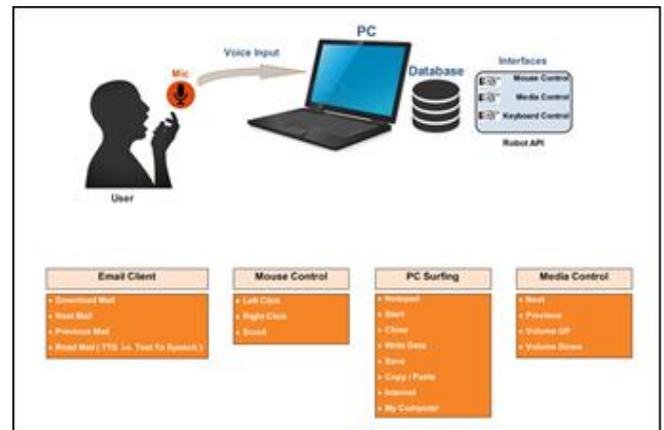


Fig. System Architecture

VI. ADVANTAGES

- The system can be used to assist dyslexic students who have problems with character or word use and manipulation in a textual form.
- The system is designed for easy accessibility of the desktop.
- The technology can help make the overall college experience more rewarding and enjoyable for individuals.
- The system uses, users command to do something via speech such as directly controlling computers, writing emails and documents, and managing the schedule.

VII. CONCLUSION AND FUTURE SCOPE

The system is designed to provide an interactive method for access the computer through speech. The system takes speech as an input and perform the respective operations on it and produces the output. This system is developed for handicapped and dyslexic people to perform their task on the computer without any difficulty. This system can be extend for the deaf peoples also who can access the system using gesture or the retina scanning technology.

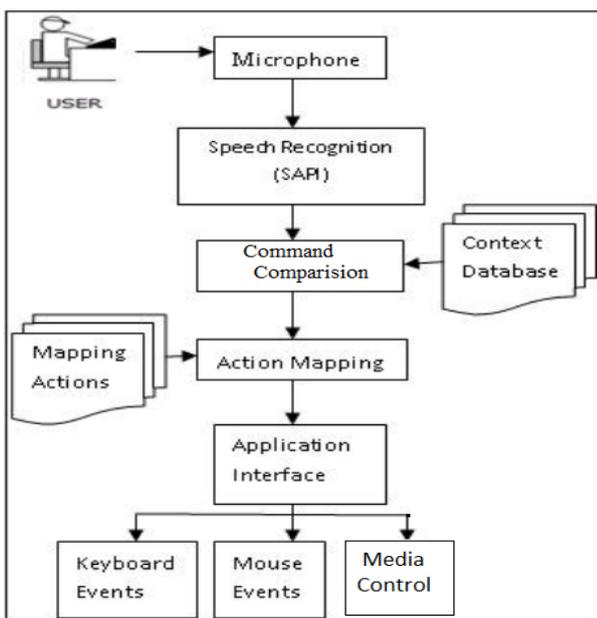


Fig. Speech Interactive System for Desktop Applications

ACKNOWLEDGEMENT

The authors would like to thanks Prof. V. S. Nandekar and anonymous referees for their constructive remarks and valuable comments.

REFERENCES

- [1] Hae-Duck J. Jeong, Sang-Kug Ye, Jiyoung Lim, Ilsun You, and WooSeok Hyun, "A Remote Computer Control System Using Speech Recognition Technologies of Mobile Devices", International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, 2012.
- [2] S. R. Patil, 2k. S. Warke, 3dhanashree Taware, 4namrata Attkare, 5divya Singh, "Speech Interactive Web Application Services", International Journal of Advanced. ISSN (PRINT): 2320-2106, 2013.
- [3] Naveenkumar C. Patil, Khyamling A. Parane, Guruprasad M. V, "4 MuteTrans: A Communication Medium for Deaf", IEEE, 2014.
- [4] Hao Shi And Alexander Maier, "Speech-Enabled Windows Application Using Microsoft SAPI" IJCSNS International Journal Of Computer Science And Network Security, VOL.6 No.9A. September 2006.
- [5] <http://docs.oracle.com/javase/7/docs/api/java/awt/Robot.html> -RobotAPI.
- [6] <http://undocumentedmatlab.com/blog/gui-automation-robot> -Robot API.
- [7] https://en.wikipedia.org/wiki/Microsoft_Speech_API - Microsoft Speech API.