

SPEECH INTERACTIVE WEB APPLICATION SERVICES

¹S. R. PATIL, ²K. S. WARKE, ³DHANASHREE TAWARE, ⁴NAMRATA ATTKARE, ⁵DIVYA SINGH

Department of Computer Engineering, Bharati Vidyapeeth's College of Engineering for Women, Katraj, Pune 411 043, India
Email: srpatil44@gmail.com, kanchan.warke@gmail.com, dhanashreetaware@gmail.com, atknamrata@gmail.com, divya.singh081@gmail.com

Abstract—Internet has become a very important factor in our day to day life. It is a wide media for communication and exchange of ideas for people staying in any nook and corner of the world. In this paper an idea is proposed to develop a speech interactive system to provide web application services. The main aim is to provide these services to the special ones who are unable to make use of the current system so efficiently. In this proposed work the main focus is on the WEB APPLICATIONS. It is tedious for the disabled people who are unable to access internet, 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. The proposed system develops the ability to handle web applications along with the O.S, mouse and keyboard control through speech. So that they can be used by disabled people without the use of the hands to develop an interface between the computer and the user. This is an attempt to develop web application through speech interaction.

Keywords- SAPI, Speech Recognition, Web Applications, Text to Speech, Speech to Text.

I. INTRODUCTION

In the recent years the technology has been drastically changing. As we know internet is playing a vital role to connect us with the outside world. It is a wide media for communication and exchange of ideas for people staying in any nook and corner of the world.

Many new developments in the technologies have been seen. So now it has become challenging to prepare a system that is unique and would provide the best services ever. It would be very boring in a situation when you are tired and you want to access your computer, and check mails. Imagine just by saying download mail the present system will download mail and also read it for you.

You need not to even enter your Email id and password and it will do it for you whatever you want through voice. And in some circumstances handicapped people will also wish to handle mouse and keyboard, email or any internet related work independently. This is the main objective of our proposed work. We are providing a speech interactive system which will work according to the user, but just by SPEECH.

As many speech enabled system are available in the market with Operating System, Keyboard, Mouse control but the proposed system will go a way beyond this and will aim to provide web applications using RSS Aggregator (Rich Site Summary). 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 [1]. It further converts this available text into speech which may be heard by the user.

II. MOTIVATION

In today's world we can get any kind of information easily with the help of internet. Web designers have the main basic aims that convey information to users or surfers in both attractive and convenient manner. However this facility can't used by handicapped people because of inefficiency of handling computer. The main motivation of the system is to address this issue.

The objective of designing interactive speech based system is to develop a framework, with the required toolset, to enable disabled people to use internet. Output should be in Speech, Text form.

III. PREVIOUS WORK

Previously various software such as Speakonia, CoolSpeech, Dragon-Naturally speaking, ReadPlease, e-speaking were designed to perform limited features as mentioned below:

- Speech Input method
- Voice Feedback
- Text Extraction and Filtering
- Text to speech

In comparison with the above software the present system will provide:

- Speech Input
- Text to Speech and Speech to Text
- RSS Aggregation for Web
- Email Handling Through Speech

- Browsing Web Content.
- Text to Braille

IV. RELATED CONCEPTS

A) SAPI

For speech synthesis and recognition Microsoft SAPI (Speech Application Programming Interface) can be used. First SAPI was introduced to Windows 95. Till now number of versions of SAPI has been released.

These versions of SAPI have been designed in a way that any software developer can create an application to perform speech recognition and synthesis with the help of a standard set of interfaces, accessible from a various programming languages. Thus we can say that SAPI is freely-redistributable component which can be shipped with any Windows application that wishes to use speech technology [1] .

In the proposed work user gives the speech or voice with the help of microphone. Microphone processes that voice to the Speech Recognition system which will convert a voice signal to a sequence of words in the form of digital data which can also be called as a command. This can be done with the help of SAPI.

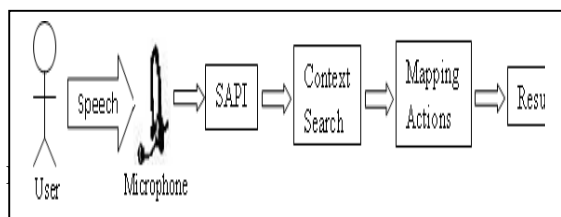


Fig: Use of SAPI

B) NanoXML Parser

XML is the extensible mark-up language, which provides a way to mark up text in a structured document. The NanoXML was first released in April 2000 as a spin-off project of the abstract user interface toolkit. It is very small and reasonably fast for xml documents and also it provides the facility of easy to use. Because of its small size, people started to use NanoXML for embedded systems (KVM, J2ME)

C) RSS Aggregator

RSS (Rich Site Summary) document is also called as “feed”. It is a family of web feed formats. It is useful to publish frequently updated works like news headlines, blogs and audio-video in standardized form. We can read this RSS feeds using software called as “RSS Aggregator” .The RSS Aggregator can be web based. RSS 0.9, the first version of RSS was created by Dan Libby and Ramanathan V. Guha at Netscape.

With the help of RSS it is possible for people that they can keep their favorite web sites in an automated manner. It is helpful to reduce task of checking them manually.

D) Text-To-Braille Converter

The Braille system is widely used by blind people to read and write. We can represent Each English alphabet in Braille character which is made-up of six dot position, arranged in a rectangle containing two columns of three dots each. For example we can define an in Braille with following combination of six dots.



Like this we can convert output i.e. text-to-Braille converter toolkit to assist a blind person to access features without the aid of a second person.

V. PROPOSED 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. The command generated by a speech recognition system is processed further for Context Search.

This process involves semantic keyword extraction and clustering from the context database. Context 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 Mouse events, Keyboard Event, O.S 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. The output can also be converted into Braille.

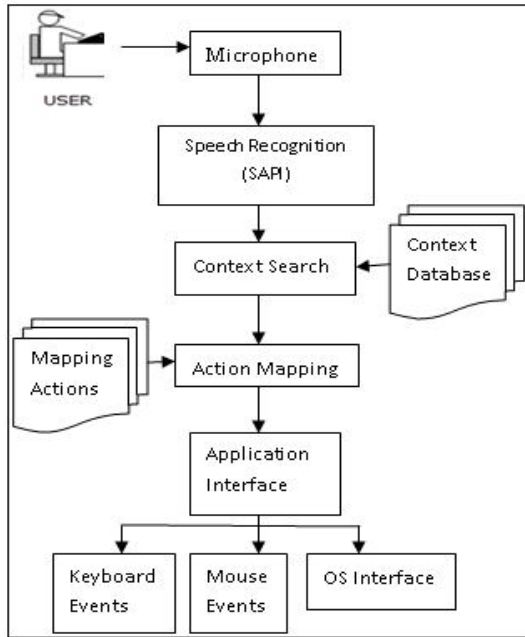


Fig: Speech interactive desktop Application

Following diagram shows the overview of a Speech Interactive Web Services. For example if we consider the email related operation.

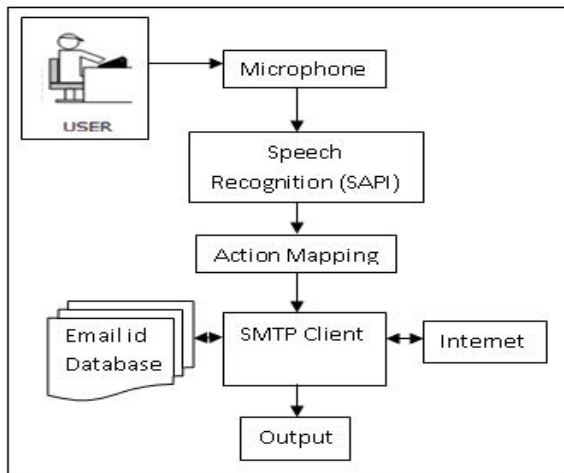


Fig: Speech Interactive Web Application for E-mail Services

Same as above user gives the speech signal with the help of microphone. 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 i.e. a command with the help of SAPI. User can operate his E-mail account with speech and without the use of the hands interfacing with commonly used human interface devices such as the mouse and keyboard. It reads out the received mails using text to speech conversion module depending upon the user command. User can also give commands to select mail, open mail, read mail, select next mail and back to home page etc. with the help of speech. We can also convert output in Braille.

Like this it is possible to provide functionality of searching, downloading and other features.

For e.g. instead of selecting by a keyboard or mouse to explore start menu we can simply say "Start" and the computer would respond by displaying start menu. By developing a fully accurate and fast speech recognition system we can possibly eliminate the use of keyboard and mouse in most of the applications[2]. Likewise disabled persons may find hands-free computing important in their everyday lives.

For accuracy purpose the context 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 (entire O.S.) can be controlled using this system.

Following diagram shows overall architecture of system.

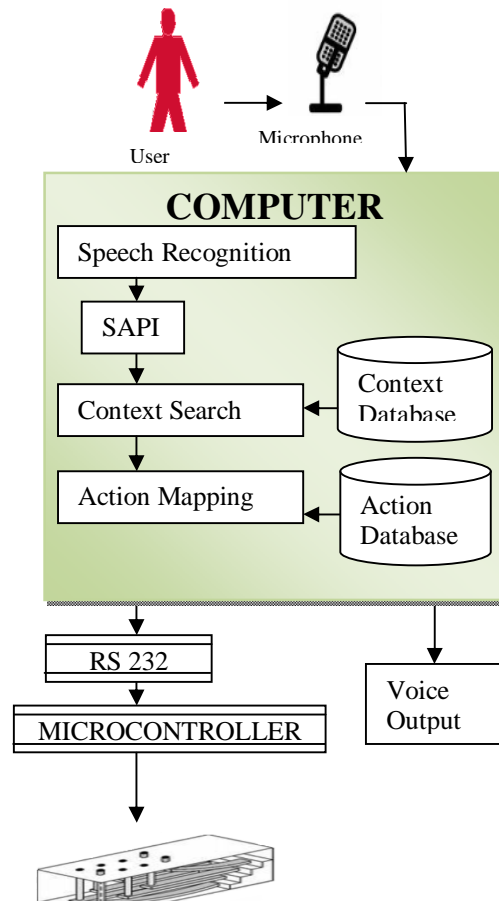


Fig: Architecture

VI. FEATURES

- 1) This system will be able to handle operating system with the help of speech (voice commands).
- 2) This system will be speech interactive means input and output is in voice.
- 3) It will give facility of speech based browser. We can navigate the web pages by giving voice commands.
- 4) Playing music just by voice commands.
- 5) This system can give facility to handle any application like notepad, Microsoft's Word just by our speech.
- 6) This system will be efficient for information retrieval and download like obtaining information about current stock, News, Weather etc.

CONCLUSION

The proposal offers a user friendly interface. This software is applicable for any version of Windows operating system. The software proposed may satisfy basic need to access operating system through their voice command which will give qualitative product for avoid time wastage and make computing so easier. It is also helpful for physically handicapped persons and visually handicapped persons for meeting their basic requirements of computing.

The intention is to help the disabled people to get the benefits of internet technologies and email facilities.

The primary aim of this project is to provide a simpler design having least cost and also reliable and practical. The system which is proposed in this paper allows the physical and visually handicapped person to operate a computer in a more efficient and natural way.

REFERENCES

- [1] 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.
- [2] Md. Abdul Kader, Biswajit Singha, And Md. Nazrul Islam., "Speech Enabled Operating System Control" IEEE 2008.
- [3] Frank Loewenich And Frederic Maire, "Motion Tracking And Specch Recognition For Hands-Free Mouse- Pointer Manipulation", ISBN 978-953-7619-29-9, PP.550, November 2008.
- [4] Dae-Young Choi, Il Kyeun Ra, "Towards A Voice Interface And Personalized Local Web Search In Smart Phones", Yuhan University, South Korea. IEEE 2010.
- [5] Ritwika Ghose, Tirthankar Dasgupta, Anupam Basu, "Architecture of A Web Browser For Visually Handicapped People", 2010 IEEE Students' Technology Symposium.
- [6] Anuj Kumar, Tim Paek, Bongshin Lee, "Voice Typing: A New Speech Interaction Model for Dictation On Touchscreen Devices", ACM 978-1-4503-1015, May 2012.
- [7] Anil J Kadam, Pallavi Deshmukh., "Speech Oriented Computer System Handling" International Conference On Intelligent Computational Systems (ICICS/2012).
- [8] Vimala. C, Dr. V. Radha., "A Review On Speech Recognition Challenges And Approaches" World Of Computer Science And Information Technology Journal (WCSIT), Vol. 2, No. 11-7, 2012.
- [9] Atiwong Suchato, Jjirasak Chirathivat, Proadpran Punyabukkana, "Enhancing a Voice Enabled Web Browser for the Visually Impaired".
- [10] Andrew Junker, Thomas, Todd Eachus, "Hands Free Computer Access For The Severely Disabled", SBIR Final Report [1R43 HD 39070-01].
- [11] <http://www.microsoft.com/speech/download/sdk>.
- [12] <http://en.wikipedia.org/wiki/XML>.

★★★