AADNYA – The Open-Office Voice Assist

¹Prof. Akshay Agarwal, ²Urvesh V. Dhangar, ³Akshay A. Dudvadkar, ⁴Pavan R. Meshram
¹Asst. Professor, ²,³,⁴BE Student, ¹,²,³,⁴Comp. Engg. Dept, SSJCET, Asangaon, India.
¹akshay1661@gmail.com, ²urvesh.dhangar11@gmail.com, ³goax2410@gmail.com,
⁴pavanmeshram21@gmail.com

Abstract - This paper is going to develop the project named "Voice Command Module for OpenOffice Writer application". The project under consideration will be a desktop application will be build around the core concept of speech recognition. Over the years HMM and GMM are very complex as they are implemented for a huge vocabulary database. Our voice command module is basically a simple module with which a user can use OpenOffice-writer application with voice commands. In our project the system will be trained by the user. When user will give voice command as input to the system it will be first checked for validation by voice module. After validation the voice command will be converts into text by voice to text conversion module. This text will be mapped with command from the database that will return the corresponding command call. Same command call will be forwarded to the OpenOffice's writer application for execution.

Keywords: *hmm algorithm, open-office, speech recognition, text to speech, vocabulary database, voice command.*

I. INTRODUCTION

This is an era of upgrade. People with limited knowledge are finding it easy to work on computers. More and more people are coming with various software which are efficient, time saving as well as effort saving. There is huge demand of such 'word processing' software products in the market such as Microsoft office.

Speech technologies are evolved to a stage where it offers great promise for human computer interaction. This speech technology can be helpful to people with minimum knowledge of computer and also for those who are physically incapable, in this circumstances software with 'command and control' facility will definitely prove useful

In this project the intent is to use speech tech with open office software which will save typing efforts and valuable time of the users while editing a document. The software comes handy when editing work is vast, complex and consists repeating tasks.

The voice command is provided as input to the application by using microphone. The corresponding command is executed in Open Office writer application as the output; hence if user wants to do any editing or formatting task he can start the application and give the desired voice command on microphone for the execution.

Voice is taken as input, it will first check for validation, and then it will convert speech to text

Converter module using SAPI .Corresponding command call will be found from the database this command call will then be forwarded to the Open-Office's writer & browser application for execution.

II. LITERATURE SURVEY

A. SAPI

Microsoft has developed Speech Application Programming Interface to allow the use of speech recognition and speech synthesis within Windows applications. A number of versions of the SAPI have been released, which has shipped either Speech SDK, or as part of Windows OS itself. In general all versions of the SAPI have been designed such that a software developer can write an application to perform speech recognition and synthesis by using a standard set of interfaces, accessible

from a variety of programming languages. SAPI versions 1 through 5 are all similar to each other, with extra features in each newer version. SAPI 5.4 is completely new interface. Broadly the SAPI can be viewed as an interface or piece of middleware which sits between applications and speech engines. In SAPI versions 1 to 4, applications could directly communicate with engines. Applications could also use simplified higher- level objects rather than directly call methods on the engines. In SAPI 5 however, application and engines do not directly communicate with other. Instead they talk to a runtime component (sapi.dll). There is an API implemented by the component which applications use, and another set of interface for engines.

The two basic types of SAPI engines are text-to-speech (TTS) systems and speech recognizers. TTS systems synthesize text strings into spoken audio using synthetic voices. Speech recognizers convert human spoken audio into readable text strings and files.

B. Openoffice.org

Openoffice.org Application commonly known Open-Office is an open-suite application whose main components are for word processing, spreadsheets, presentations, graphics, database etc. It is distributed as free software and is written using its own GUI toolkit. It supports ISO/IEC standard Open Document Format (ODF) for data interchange as its default file format, as well as Microsoft Office formats among others. As free software, users are free to download, modify, use and distribute OPenPffice.org and its components.

1. Writer: A word processor similar to Microsoft Word. It can export portable document format (PDF) files, and can function as a basic WYSIWYG editor for creating and editing web pages
2. Calc: A spreadsheet similar to Microsoft Excel. Calc can export spreadsheets to the PDF format. Calc provides a number of features not present in excel.
3. Base: A database management program similar Microsoft Access. IT allows creation and manipulation of databases, and the building of forms and reports to provide easy access to data for end user.
4. Draw: A vector graphics editor comparable in features to early versions of CorelDraw and Microsoft Visio. It features versatile "connectors" between shapes, which are available in a range of line styles and facilitate building drawings such as flowcharts
5. Math: A tool for creating and editing mathematical formulae.
6. Impress: A presentation program similar to Microsoft PowerPoint. Impress can export presentations to Adobe Flash (SWF) files, allowing them to be played on any computer.

III. AIM AND OBJECTIVE

Speech recognition forms the core of the project, the application tries to target the needs of today's time that are time saving, less efforts, and ease to access. Aadnya - The Open-Office Voice Assist is software that will enable command and control to Open-Office application.

Objective

1. People with disability: This application can be of great use to the visually and physically challenged people especially for people who have difficulty using their hands, involved disabilities that preclude using conventional computer input devices. Also for people who used the keyboard a lot and developed RSI will find this software helpful. In fact this software will prove to be a boon for them.

2. Education: Due to simplicity of this software, it can prove helpful for teaching and training office productivity software (in this case OpenOffice.org) to a new-b users.

3. Schools and colleges: It can be also used in schools and colleges for learning and study purpose. It can be used in schools for students with disabilities.

4. Offices: In offices this software utility can prove very helpful for saving time and efforts for documentation and formatting purpose. In fact this software will bring increased productivity.

IV. STAGES IN PROPOSED SYSTEM

1. Voice will be taken as input from the microphone having good quality of noise reduction.
2. System will pass this voice input to Microsoft Speech API which will convert speech into text.
3. This text will then be processed by mapping module to map to corresponding command (Mapping module will

have all the text command and their corresponding shortcut keys mapped).

4. Then the command found will be passed to running instance of open office application.

5. Then the command will be executed.

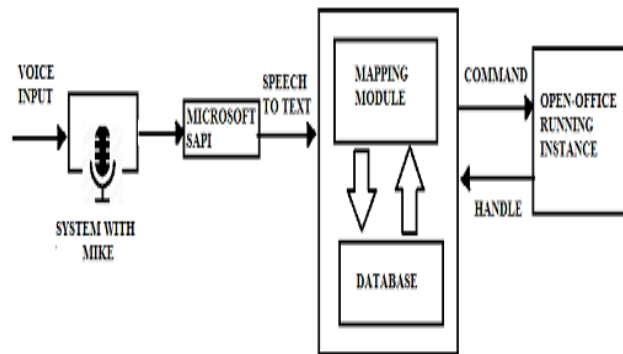Above steps can be shown as follows:

Fig 1: Proposed System

The OpenOffice voice assist is software that enables command and control to OpenOffice application. The input provided to the application will be voice command like select all, cut, copy, paste etc. that will be converted into speech to text by speech to text conversion module. After the get command on text format it will be checked for its presence in commands available for execution. Once the command found in database corresponding command call will be taken to the OpenOffice application. This command to command call mapping will be done by mapping module. The system will be built around the core e concepts of speech recognition which convert the spoken words into text; which makes communication between humans and computer easier and convenient.

Following are some of the commands in Open-Office with their shortcut keys:-

Table 1: Commands and their shortcut keys

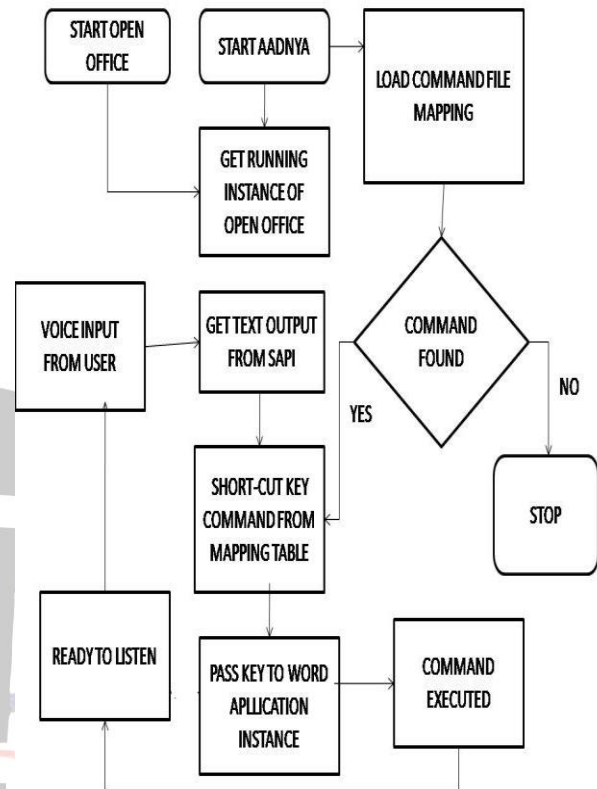| Commands	Shortcut keys
Copy	Ctrl+c
Paste	Ctrl+v
Cut	Ctrl+x
Select all	Ctrl+a
Delete	Del
Find	Ctrl+f
Find next	F3
Replace	Ctrl+h

V. ALGORITHMS IMPLEMENTATION

Fig 2: Voice Command Algorithm

VI. EXPECTED OUTPUT

The user gives the voice command to edit or modify the file which is currently running. The voice command is converted internally into text and is searched for its short-cut key in the command mapping module, and the command is executed.
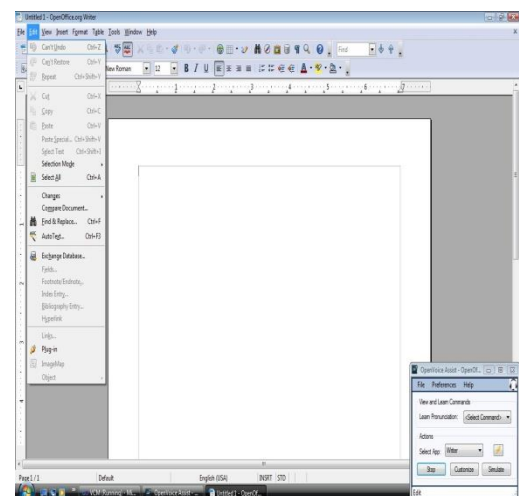

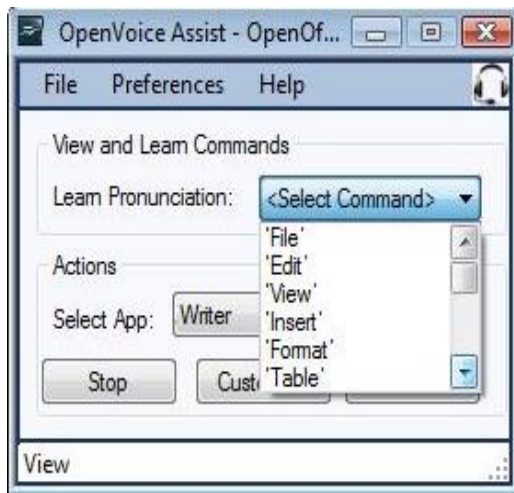**Fig 3:Running instance of open-office**

Fig 4:OpenVoice Assist GUI

VII. CONCLUSIONS

Even to solve issue of pronunciation have provided the facility of 'Learn Pronunciation'. It will tell user how to correctly pronounce the command selected. The simulate facility is provided to make user aware of available commands for execution.

Such is the utility of the software that it makes daily work in offices or at homes very easy. Thus such software in itself is very leisurely and technically apt and hand in hand with developing technology. This software can prove useful to visually and physically challenged people as well.

Moreover this software not only performs speech recognition but also gives response to user in event of accurate speech recognition. When it recognizes a voice command successfully, it performs speech synthesis using Microsoft text-to-speech conversion facility and notifies user accordingly. This functionality of the software proves helpful for visually challenged user. For New-B user the given facility of simulation of software or demo. This application can be used in all those domains where time and effort saving is preferred for text formatting and documentation purpose giving enhanced user experience. This developed modular application such that by making few changes this application can later be used to automate any application like Mozilla Firefox.

REFERENCES

[1] Alwang, Geg. "*Speech Recognition*", PC Magazine, December 1, 1999.

[2] Hauptmann, Alexander G. Jang, PhotinaJaeyun. Caenegie Mellon University. "*Learning to Recognize Speech by Watching Television*, "IEEE Intelligent System, and September/October 1999.

[3] Miastkowski, Stan. "*Latest Speech Software Gets You Up and Running Faster*, "PC World, November 1999.

[4] Spoken Language Processing*: A Guide to Theory, Algorithm and System Development* Xuedong Huang, Alex Acero, Hsiao-Wuen Hon

[5] Lawrence R. Rabiner, *A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*. Proceedings of the IEEE, 77 (2), p. 257286, February 1989.

[6] V. Petrushin. *Hidden Markov Models: Fundamentals and Applications. Part 2: Discrete and Continuous Hidden Markov Models, Online Symposium for Electronics Engineers. 2000*

[7] King, A., Evans, G., &Blenkhon, P., WebblE*: a web browser for visually impaired people*. In S. Keates, P.; Clarkson, P. Langdon, & P. Robinson (Eds), Proceedings of the 2 Cambridge Workshop on Universal Access and Assistive Technology, Springer-Verlag, London, UK, 2004

[8] Ghose R, Dassgupta T, Basu A, *"Architecture of a web browser for visually handicapped people*", students technology symposium (TechSys),2010 IEEE.

[9] Manaris, B. and MacGyvers, V. *"Speech for Alphanumeric Data Entry - A Usability Study,"* submitted to the International Jthisnal of Speech Technology.

[10] Harkreader and Manaris, B, A. SUITE: speech understanding interface tools and environments, in Proceedings of FLAIRS '97 (Daytona Beach, FL, May 1997), 247-252.Ben Mosbah, B., "*Speech Recognition for Disabilities People*" Volume 1,Information and Communication Technologies, 2006. ICTTA apos;06.2nd,Issue , 24-28 April 2006 Page(s): 864 – 869

[11] Mukund Pabmanabhan, Michel Pichney, "*Large vocabulary speech recognition algorithms*", IEEE Computer magazine, 0018-9162/02, pp. 42-50, 2002.

[12] Md. Abdul Kader, Biswajit Singha, and Md. Nazrul Islam, "*Speech Enabled Operating System Control*" Proceedings of 11th International Conference on Computer and Information Technology (ICCIT 2008) 25-27 December, 2008, Khulna, Bangladesh 1-4244-2136-7/08 IEEE 2008

[13] D. LeBlanc, Y. Ben Ahmed, S. Selouani, Y. Bouslimani, H. Hamam, "*Computer Interface by Gesture and Voice for Users with Special Needs*" 1-4244-0674-9/06 IEEE 2006.