

JARVIS: Voice Controlled AI for Help Human

¹Prof. Rupali Panpaliya, ²Ashish Gaikwad, ³Prathmesh Kamble, ⁴Rushikesh Kad, ⁵Saurabh Pasalkar

^{1,2,3,4,5}Department of Computer Science and Engineering, Siddhant College of Engineering, Sudumbare, Pune, India

Authors E-mail: ²ashishgaikwad10@gmail.com, ³deshmukhprathmesh25@gmail.com, ⁴kadrushi712@gmail.com, ⁵pasalkarsaurabh935@gmail.com

Abstract - This Paper considers an overview of speech recognition technology, Software development, and its applications. The first section deals with the description of speech recognition process, its applications in different sectors, its flaws and finally the future of technology. Later part of report covers the speech recognition process, and the code for the software and it is working. Speech Recognition is the process of automatically recognizing a certain word spoken by a particular speaker based on individual information included in speech waves. In this project, we will use algorithms for the speech recognition which will implement on JAVA for platform independent facility this system can be used for any security system in which the person authentication is required.

Keywords: Speech recognition, signal processing.

1. Introduction

Speech Recognition is the process of automatically recognizing a certain word spoken by a particular speaker based on individual information included in speech waves. This technique makes it possible to use the speaker's voice to verify his/her identity and provide controlled access to services like voice based biometrics, database access services, voice based dialing, voice mail and remote access to computers. In this project, we will use algorithms for the speech recognition which will implement on JAVA for platform independent facility. This system can be used for any security system in which the person authentication is required.

Speech is the most natural way to communicate for humans. While this has been true since the dawn of civilization, the invention and widespread use of the telephone, audio-phonetic storage media, radio, and television has given even further importance to speech communication and speech processing.

The advances in digital signal processing technology has led the use of speech processing in many different application areas like speech compression, enhancement, synthesis, and recognition humans. While this has been true since the dawn of civilization, the invention and widespread use of the telephone, audio-phonetic storage media, radio, and television

has given even further importance to speech communication and speech processing.

The objective of the current research is

- To understand the speech recognition and its fundamentals
- Its working and applications in different areas
- Its implementation as a desktop Application
- Development for software that can mainly be used for:

- 1) Speech Recognition
- 2) Speech Generation
- 3) Text Editing Tool for operating Machine through voice.

2. Problem Statement

In this 20th century everything is based on Computer. As we can see there are lots of enterprises which are developing so advanced technologies that can true shape the future of our world that we are living in. Nowadays we use several different technologies that make our life easy.

So we required various software for making our life quite a bit easy. We required software for protection and security of our machine, developers use eclipse for making programming easy, we use Google engine for searching etc. In short we are surrounded by software.

3. Proposed System

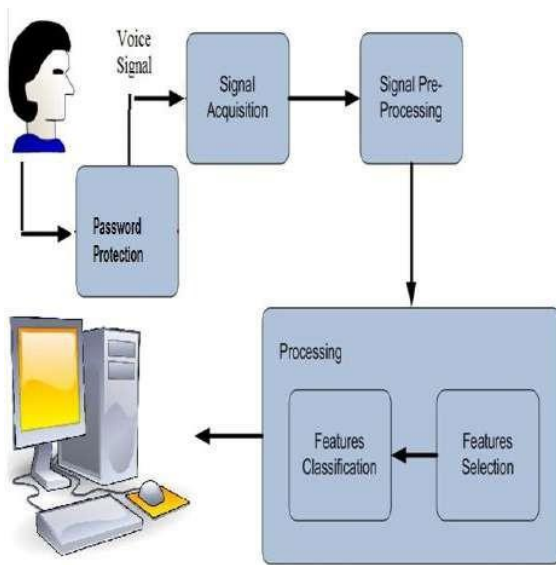
In Today's world we use keyboard and mouse for operating our computer normally. This is where our software comes into practice. Using this software we can fully operate our computer with our own voice. Its sounds a quite hi-tech and quite impossible but it's true, this can happen.

Today various company our working on it to make advance into this Speech Recognition System technology. Cause the system will fully reduce the use of the keyboard and just. Just you will require a Mic or Bluetooth device from which we can give a voice input to the system.

It's all in one technology as for security mechanism you will use your voice as a security password you will operate the computer with our voice and perform the entire task that do by

using keyboard and mouse. Besides this you need to sit in front of our computer for doing the task, use SRS you can perform this entire task by sitting in your room and relax.

So we required various software for making our life quite a bit easy. We required software for protection and security of our machine, developers use eclipse for making programming easy, we use Google engine for searching etc. In short we are surrounded by software.



4. System Architecture

Module 1: Speech Command Operations

It's our project most important module. It contains various commands to perform regular computer operations and functions. The speech is recognized from the mic and then compared with the conditions given and then the operations are performed. In this module we have made use of wired headset for taking the voice input. The module contains a dict. gram file which contains various commands like computer, browser, paint, run, copy, close, delete, paste, and space, open which we use to run the functionalities in computer like My Computer, MS-Paint, Browser, Face book, G-mail etc.

Module 2: Remote Based Innovation

In this module we have made use of Bluetooth headset for taking the voice input and giving it to the system for performing task. The main advantage of this module that we can operate our computer within 5 meters of range of the computer. It provides user a greater comfort as they can operate the computer from a longer distance doing their other task simultaneously. As we all know that life in the present world has become very much fast and hi tech, so we decided to make a project related to it.

Module 3: Speech to Text

In this module we have taken the voice input from the Mic and printed that on the notepad. In addition to that we can also perform various task on that such as copy, paste, select all, delete etc. As we all know that detection of speech and converting that into the text format is quite a complicated job so in order to make it a bit less complex we have made use of a file which contains various keywords which we will be going to use in our application. The module works by taking the input from the user, it then compares the given input with the file, if the keywords given as input by the user matches with the words or commands mentioned into the file then it prints those words onto the notepad or word document.

5. System Design

1) Working

This software named 'J.A.R.V.I.S' has the capability to write Spoken words into text are of notepad, and also can recognize your Commands as "save, open, clear" this software is capable of opening Windows software such as notepad, MS-Paint, calculator through voice Input. This actual working of the system can represent educing the diagram Given below

Opening My Computer

Software is able to open the My Computer by just giving the giving The command computer.

Opening Command Prompt

Software is able to open the Command Prompt by just giving the Command run command prompt.

Opening Browser

Software is able to open the Google by just giving the command open Browser. It can also open sites directly like Facebook, Gmail, etc. just we need is to speak their commands after opening the browser.

6. Technical Requirements

Hardware Requirements: Processor Intel Pentium IV or later
Ram Minimum 512 MB Secondary storage Minimum 80 GB

Software Requirements: Operating System Server Windows XP or later Database Server Oracle or MYSQL Client Microsoft Internet Explorer Tools and environment Apache Tomcat 1.6, jdk1.6 User Interface Html 5, Jsp Code Behind Java, Jdbc, MySQL 5.5, Servlet.

7. Conclusion

As we all know that life in the present world has become very much fast and hi-tech, so we decided to make a project related to it. Nowadays, people feel tired of using the computer as we have to give the commands to it using the keyboard or mouse and it has now become a little bit of old generation, so we decided to make a project that will change this old and low tech method. Our project is fully based upon the concept of voice recognition activation. Imagine a computer that will be fully controlled with your voice, you just have to speak a task and it will be performed by the computer, besides this we have also added the feature of remote access which allows you to access your computer by sitting at the home. This project is also useful for the peoples who are physically impaired.

We have built this project for the following needs of the mankind.

- 1) Hi-Tech Technology Experience
- 2) Access for physically impaired people
- 3) Security Management Purpose
- 4) Time Saving Technology
- 5) Remote Access Control Mechanism.

8. Future Scope

So we required various software for making our life quite a bit easy. We required software for protection and security of our machine, developers use eclipse for making programming easy, we use Google engine for searching etc. In short we are surrounded by software.

There are currently many companies which are constantly making progress in this SRS technology. Their view is that the computer should be operated with more ease and by everyone. Besides these it also makes humans life much easier. Because we just need to speak and the computer will be operated.

In advancement to that it has been heard that the Microsoft Company will soon be developing an Operating System that will be operated with our voice only. It will totally demolish the use of keyboard and mouse forever.

ACKNOWLEDGEMENT

We would like to take this opportunity to thank all the people who were part of this seminar in numerous ways, people who gave un-ending support right from the initial stage.

In particular we wish to thank Prof. Rupali as internal project guide who gave their co-operation timely and precious guidance without which this project would not have been a

success. We thank them for reviewing the entire project with pains taking efforts and more of her, unbanning ability to spot the mistakes.

We would like to thank our H.O.D of Computer Department Prof. Sushma Shinde for her continuous encouragement, support and guidance at each and every stage of project.

REFERENCES

- [1] L.J"ic~'eck, I. Maicssc, M. Ga~ si~ c, S. Kcizci, B. l'kom-~ so, K. Y",and S. Young, "Tiansfoimatiobased Lcaiqiig roi semantic paising," inNioceediigs or INTERSNEECH, 2009.
- [2] Iiaccesco Nesta, Maico Matasso, Ramo Icaadcz Ast"dillo, "A Ilcxible Spatial Bliid So"icc Extfatio Iamcwoik Ioí Rob"st Spceck Rccogitio I NoisQ Enviionments", Comp"teí Spceck a d La"g"agc.
- [3] H. S"z"ki, H. Zc, Y. N"ok"k", C. MiQajima, K. lok"da, a d I.Kitumuíu, "Speech Recognition Using Voice Ckafacticístic Kcpc d c t Aco"stic Models",Acoustics, Spceck, a d Sig al Píocessi g, 200«. Píocccdi gs, MaQ 200«.
- [4] Ji mook Lcc, Sco gwook Paik, I joo Ho g a d HoiJ" Yoo, "An E cigQ- crricic t Spceck Extfatio Píocessoí roi Rob"st Uscí Spceck Rccogitio i Mobile Head-mo" tcd KisplaQ Systems", IEEE ía sactio s o Cíc"its a d SQstems II, Vol"mc: 64, Iss"c: 4, Apíl 2017.
- [5] Mickacl Píicc, Me½beí, James Glass and Anantha P. Chandiakasan, "A Low- Powcí Spceck Rccogizcí a d Voicc Acti:itQ Kctctoí Usi g KccpNe"íal Networks", IEEE Jo"íal or Solid-Statc Cíc"its, Octobcí 2017.
- [6] Ng"Qc K"c l'ka g, S" gQo" g Lcc, Yo" g-Koo Lcc, KQ" g Hcc, "Fast Co stíai cd I d cpc d c t Compo c t A alQsis Ioí Bliid Spceck Scpafatio Witk M"ltiple Refeénces",Computeí Scic ccs a d Co :cíg cc I íoimatio l'ck ologQ (ICCI) Kcc. 2010.
- [7] Pkilipp Aicki gci, Maíti Hagm"llcí, Bcít Sck cidcíSticklci, Jca Schoentgen and Fíanz Peínkopf, "Tíacking Of Multiple Fundamental Íicq" cics i Diplophonic Voices", IEEE/ACM íiansactions on Audio, Spceck, a d La"g"agc Píocccsi g, Vol"mc: 26, Iss"c: 2, Octobcí 2017.
- [8] Suma Swamy and K.V Ramakiishnan, "An Efficient Speech Recognition System", Computeí Science & Engineeíng: An I tci atio al Jo"íal (CSEIJ), Vol. «, No. 4, A" g"st 201«.
- [9] IEEE Website: <http://www.ieee.com>

- [10] Google website: <http://code.google.com/SRS>
- [11] <http://www.wikipedia.com/Speech>
- [12] <http://www.google.co.in>

- [13] Referred Books: For Core JAVA, “JAVA complete reference” by Herbert Schildt.
- [14] For Speech Recognition, “Speaker Identification” by Lawrence R.

Citation of this Article:

Prof. Rupali Panpaliya, Ashish Gaikwad, Prathmesh Kamble, Rushikesh Kad, Saurabh Pasalkar, “JARVIS: Voice Controlled AI for Help Human” in proceeding of International Conference of Recent Trends in Engineering & Technology ICRTET - 2023, Organized by SCOE, Sudumbare, Pune, India, Published in IRJIET, Volume 7, Special issue of ICRTET-2023, pp 157-160, June 2023.
