International Journal of Advanced Trends in Computer Applications
www.ijatca.com

Automatic Speech Recognition through Artificial Intelligence

[1]Dr. M. Thangamani, [2]Ms. P. Seetha Subha Priya
[1, 2]Assistant Professor, Kongu Engineering College, Perundurai-638052
[1]manithangamani2@gmail.com, [2]sspriyam@gmail.com

Abstract: *The field of computer science deals with designing computer systems that can recognize spoken words. Note that voice recognition implies only that the computer can take dictation, not that it understands what is being said. Comprehending human languages falls under a different field of computer science called natural language processing. A number of voice recognition systems are available on the market. The most powerful can recognize thousands of words. However, it generally requires an extended training session during which the computer system becomes accustomed to a particular voice and accent. This proposed system control the system operations through voice.*

Keywords: *Speech recognition, Artificial network, recognition*

1. Introduction

The user communicates with the application through the appropriate input device i.e. a microphone. The Recognizer converts the analog signal into digital signal for the speech processing. A stream of text is generated after the processing. This source-language text becomes input to the Translation Engine, which converts it to the target language text.

A number of voice recognition systems are available. But they commonly need an extended training session during which the computer system becomes adapted to a particular voice and enunciation. Such systems are said to be speaker dependent. Many systems also require that the speaker speak slowly and distinctly and separate each word with a short pause. These systems are called discrete speech systems. Recently, great strides have been made in continuous speech systems. Voice recognition systems that allow you to speak naturally. There are now several continuous-speech systems available for personal computers.

Because of their limitations and high cost, voice recognition systems have traditionally been used only in a few specialized situations. For example**,** such systems are useful in instances when the user is unable to use a keyboard to enter data because his or her hands are occupied or disabled. Instead of typing commands the user can simply speak into a headset. Increasingly, however, as the cost decreases and performance improves, speech recognition systems are entering the mainstream and are being used as an alternative to keyboards.

2. Existing System

In the existing system, the user should well know to handle the system and describe the applications with proper manner. Educated persons only use the computer and search the specified content or folder or programs or some others etc. lot of time waste in the previous system. Have to search the specific by typing the details using the keyboard. Manual work is increased and there is no proper solution for the searching.

3. Related Research

Speech recognition technology has made it possible for computer to follow human voice commands and understand human languages [1]. The main goal of speech recognition area is to develop techniques and systems for speech input to machine speech is the most natural form of human communication. The probabilistic framework, describes how to represent and manipulate uncertainty about models and predictions, has a central role in scientific data analysis, machine learning, robotics, cognitive science and artificial intelligence [2]. They discussed some of the state-of-the-art advances in the field, namely, probabilistic programming, Bayesian optimization, data compression and automatic model discovery.

Many speech recognition systems are used Isolated word speech recognition systems require that speaker must paused briefly between the words while continuous speech recognized system does not spontaneous speech may have disfluesncies and is difficult to recognize[3, 4, 5]. Speech recognition or speech to text includes capturing and digitizing the sound waves, transformation of basic linguistic units or phonemes, constructing words from phonemes and contextually analyzing the words to ensure the correct spelling of words that sounds the same [6]. They discovered neural network based solutions of speech recognition tasks, detecting signals using angular modulation and detection of modulated techniques.

Voice control system for the use in the robotized manufacturing cells as well as to create tools providing its simple integration into manufacturing [7]. The algorithm for semantic analysis, using specific features of voice commands used for controlling industrial devices and machines, has been incorporated into the system. Voice activated command and control framework for the control of remote devices in a ubiquitous computing environment is exposed [8]. The prototype device is a Java controlled Lego Mindstorm robot. The research considers three different scenario configurations. A recognition grammar for command and control of the robot has been created and implemented in Java, in part in the recognition engine and in part on the robot.

Fuzzy modeling require two main steps which are structure identification and parameter optimization, the first one determines the numbers of membership functions and fuzzy if-then rules, while the second identifies a feasible set of parameters under the given structure. However, the increase of input dimension, rule numbers will have an exponential growth and there will cause problem of “rule disaster”. Samiya Silarbi et. al [9] applied adaptive network fuzzy inference system ANFIS for phonemes recognition.

4. Problem Definition

The speech quality varies from person to person. It is therefore difficult to build an electronic system that recognizes everyone’s voice. By limiting the system to the voice of a single person, the system becomes not only simpler but also more reliable. The computer must be trained to the voice of that particular individual. Such a system is called speaker dependent system.

Speaker independent systems can be used by anybody, and can recognize any voice, even though the characteristics vary widely from one speaker to another. Most of these systems are costly and complex. Also, these have very limited vocabularies. It is important to consider the environment in which the speech recognition system has to work. The grammar used by the speaker and accepted by the system, noise level, noise type, position of the microphone, and speed and manner of the user’s speech are some factors that may affect the quality of speech recognition.

5. **Proposed Architecture**

This work controls the system operations through voice. Voice is given as input to activate any system commands such as Executable File Open, Shutdown and Reset. This work is designed with Speech SDK software and Microphone. When the user’s voice is fed into the system through microphone, the SDK software converts the voice into its equivalent string format. If the voice command is fed as “Word” in the microphone, this voice format is converted into string format and it is compared with original input voice and the Microsoft Word will open. In the same way the other File like Games, CD Drive Open and Close, Notepad Open and Microsoft-PowerPoint etc.

Similarly, the system recognizes the user’s voice as input data and interprets it into text matter on the screen. This operation of the system reduces the workload of the user in entering the data through keyboard. If you want to Save the Documents this option is also available in this project. The file is saved as notepad file.

By using this project we can control the system and applications through the speech signal. We have designed the Visual basic.Net as Front end Application. Visual Basic is referred to as an “Event Driven” programming language because all the code is triggered by specific events that the user performs.

The user communicates with the application through the appropriate input device i.e. a microphone. The Recognizer converts the analog signal into digital signal for the speech processing. A stream of text is generated after the processing. This source-language text becomes input to the Translation Engine, which converts it to the target language text.

Voice or speech recognition is the ability of a machine or program to receive and interpret dictation, or to understand and carry out spoken commands. For use with computers, analog audio must be converted into digital signals. This requires analog to digital conversion. For a computer to decipher the signal, it must have a digital database, or vocabulary, of words or syllables, and a speedy means of comparing this data with signals. The speech patterns are stored on the hard rive and loaded into memory when the program is run. A comparator checks these stored patterns against the output of the A/D converter. Figure 1 shows the Voice processing architecture for proposed system.

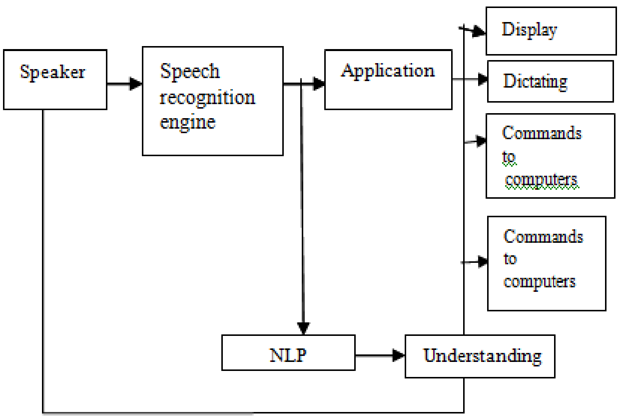

Figure 1. Proposed Architecture of Voice processing

In practice, the size of a voice-recognition program's effective vocabulary is directly related to the random access memory capacity of the computer in which it is installed. A voice-recognition program runs many times faster if the entire vocabulary can be loaded into RAM, as compared with searching the hard drive for some of the matches. Processing speed is critical as well, because it affects how fast the computer can search the RAM for matches. All voice-recognition systems or programs make errors. Screaming children, barking dogs, and loud external conversations can produce false input. Much of this can be avoided only by using the system in a quiet room. There is also a problem with words that sound alike but are spelled differently and have different meanings -- for example, "hear" and "here." This problem might someday be largely overcome using stored contextual information.

6. Discussion

In the proposed system, the system is more advanced of searching the specified content. Lot time saved in this process and no need to search the content using the keyboard. This system is based on speakers, microphones and sound cards. Based on the voice command the system will automatically activate and do the process by intimating commands. For example command give like start, then the start window opens automatically. Figure 2 and 3 shows the command give as control panel, then the control panel window opens automatically.

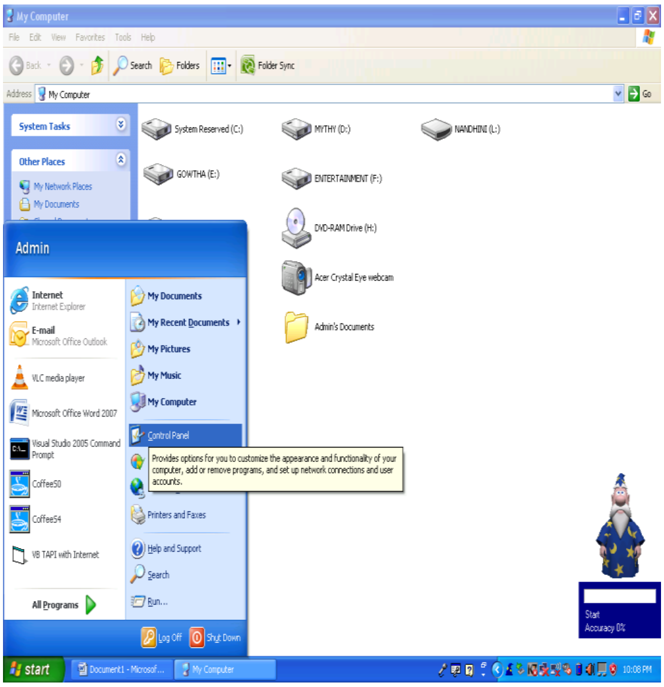

Figure 2 Opening control panel window

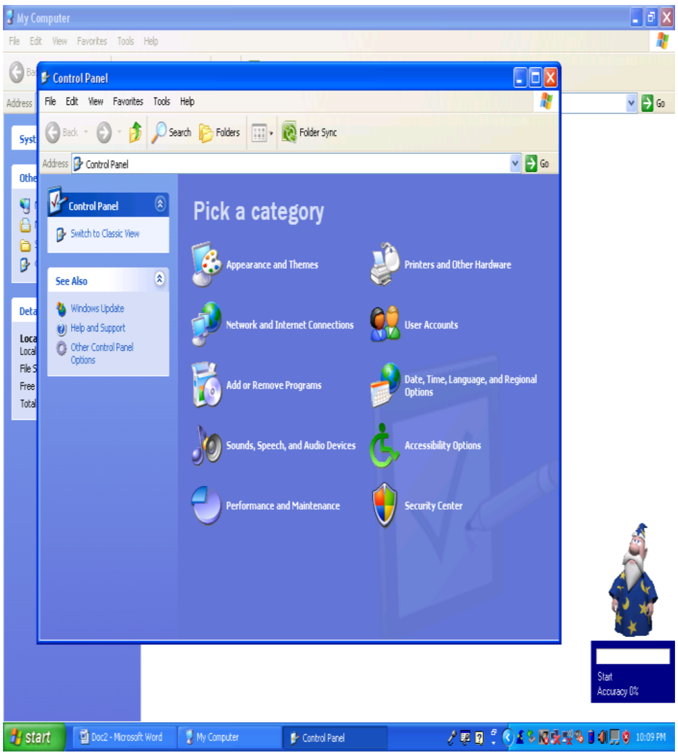

Figure 3 Opening control panel window final screen

7. Conclusion and Future work

In this research we created a generation of devices that provide a rich interaction experience. For situations where a visual human machine interface is unsuitable or too expensive, voice is a practical alternative. Voice output has been available for some time, but voice control has been a unique challenge even though there are a number of embedded applications. The challenge with voice control is overcoming the variance between human voices, which can be more or less of an issue depending on the desired implementation. As a result of changes in shape of human vocal tract during generation of different words, resonance frequencies of vocal tract, formants, also changes. Using this phenomenon, we can extract voice features of each command and we can implement a voice command recognition system. In training phase, if stated voice commands contain more vowel differences between them, we will have more accurate recognition system. Accuracy of system also increases if we increase number of repetitions for each command in training stage.

References

[1] Himanshu, Sarbjit Kaur, Literature Survey on Automatic Speech Recognition System, Vol 4, No. 7, 2014 .
[2] Zoubin Ghahramani, Probabilistic machine learning and artificial intelligence, Vol.521, 452–459, 2015.
[3] X. D. Huang, Y. Ariki, M.A. Jack Hidden Markov models for speech Recognition, The Complete Practical Reference Guide; T. Schalk, P. J. Foster: Telecom Library Inc, New York
[4] C.H. Lee, F.K. Soong and K.K. Paliwal , ", Kluwer, Boston, Automatic Speech and Speaker Recognition: Advanced Topics1996.
[5] S. E. Levinson, L. R. Rabiner and M. M. Sondhi : An Introduction to the Application of the Theory of Probabilistic Functions of a Markov Process to Automatic Speech Recognition in Bell Syst. Tech. Vol. 62, No. 4, pp. 1035-1074, 1983.
[6] Khaldoon A. Ghaidan, Huthaifa A. Issa, Esam Trad, Khalid Smadi, Artificial Intelligence for Speech Recognition Based on Neural Networks, Journal of Signal and Information Processing, Vol. 6, pp. 61-72, 2015.
[7] Adam Rogowski, Industrially oriented voice control system, Robotics and Computer-Integrated Manufacturing , Elsevier , Vol.28, pp.303–315, 2012.
[8] Tony Ayres, Brian Nolan, Voice activated command and control with speech recognition over WiFi, science of Computer Programming, Elsevier, Vol. 59 pp.109–126, 2006.
[9] Samiya Silarbi, Bendahmane Abderrahmane,and Abdelkader Benyettou, Adaptive Network Based Fuzzy Inference System For Speech Recognition Through Subtractive Clustering, International Journal of Artificial Intelligence & Applications (IJAIA), Vol. 5, No. 6, pp.43-51, November 2014

Authors Biography

Dr. M. Thangamani is nearly 20 years of experience in research, teaching, consulting and practical application development to solve real-world business problems using analytics. Her research expertise covers data mining, machine learning, cloud computing, big data, fuzzy, soft computing, ontology development, web services and open source software. She has published 50 articles in International journals and presented over 67 papers in national and international conferences in above field. She has delivered more than 35 Guest Lectures in reputed engineering colleges on various topics. She has organized many self supporting and sponsored national conference and Workshop in the field of data mining, big data and cloud computing. She is on the editorial board and reviewing committee of leading research journals, and on the program committee of top international data mining and soft computing conferences in various countries. She also seasonal reviewer in IEEE Transaction on Fuzzy System, international journal of advances in Fuzzy System and Applied mathematics and information journals. She has organizing chair and keynote speaker in international conferences in India and abroad. She is currently working as Assistant Professor in Kongu Engineering College.

Ms. P.Seetha Subha Priya has completed Master of Computer Application. Her research expertise covers data mining, Speech recognition, cloud computing and big data. She is currently working as Assistant Professor in Kongu Engineering College.