

International Journal of Advance Research in Computer Science and Management Studies

Research Article / Survey Paper / Case Study

Available online at: www.ijarcsms.com

A Survey of Voice Recognition System and its Application

Rutuja Balasaheb Dhobale¹

Department of Computer Engineering,
Anantrao Pawar College of Engineering and Research
Pune, India

Manoj Mulik²

Department of Computer Engineering,
Anantrao Pawar College of Engineering and Research
Pune, India

Abstract: Voice recognition is process of converted voice into electrical signals, and these signal transformed into text. It is also known as "automatic speech recognition" (ASR), "computer speech recognition", or just "speech to text" (STT). . There are many use of voice recognition like Lecture translates in the text format. Voice recognition have advantages like reduce the work, helps to people which can't hear and also blind people, security. The important factors are considered in the Voice recognition like quality of hardware like microphone; take care of bad voice, some distance between two words', grammar

Keywords: Voice recognition, Grammar, Automatic Speech Recognition.

I. INTRODUCTION

Voice recognition is very interesting topic in computer field. In this topic we study how to communicate with system easily using voice. Voice recognition system simply defines, "software allows users to translate their spoken words into written words through the use of microphones". Voice recognition used in real world like perform operation on computer using voice command, used the voice compiler to convert one language to other. In advanced we used in the mobile, car etc. In this the voice command are given to hardware device it perform processing and appropriate action is perform.

Speech Engine: The Speech Engine loads a list of words to be recognized. This list of words is called a grammar. Takes input as distinct characteristics of sound - derived from the waveform and compares them with its own acoustic model. The engine searches its acoustic space, using the grammar to guide this search. It then determines which words in the grammar the audio most closely matches and returns a result.

In the speech engine component

1. Human User: The user enters the commands.
2. Voice Recognition: Recognize the voice create waveforms and convert into text
3. Machine: In this machine the text which produced by the voice recognition system it check the grammar.

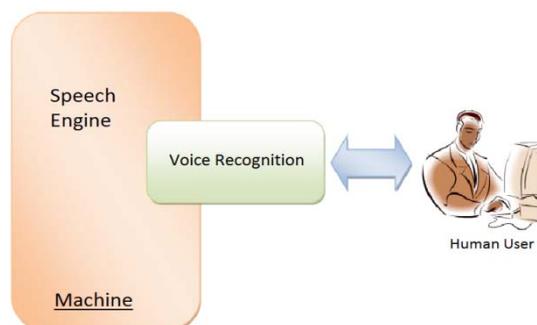


Fig.1: Speech Engine.

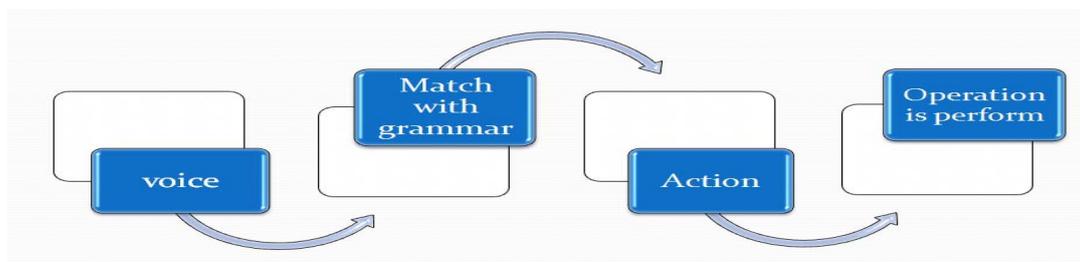
Basic Steps of Operation Using Voice Command:

Fig. 1 Basic Steps of Operation Using Voice Command

1. Voice: In that the voice is accepted convert into text.
2. Match with Grammar: in that phase the check the command properly given or not. In that the accepted word /command matches with the dictionary.
3. Action: when the word is match with dictionary word it performs the action (execute the small program of command).
4. Operation Is Perform: In this the commands is executed and give the result.

Voice Recognition: Voice recognition provides computers with the ability to listen to spoken language and determine what has been said. In other words, it processes audio input containing speech by converting it to text.

The major steps of a typical speech recognizer are as follows:

1. Grammar design: Defines the words that may be spoken by a user and the patterns in which they may be spoken.
2. Signal processing: Analyze the spectrum (the frequency) characteristics of the incoming audio.
3. Phoneme recognition: Compares the spectrum patterns to the patterns of the phonemes of the language being recognized.
4. Word recognition: Compares the sequence of likely phonemes against the words and patterns of words specified by the active grammars.
5. Result generation: Provides the application with information about the words the recognizer has detected in the incoming audio.

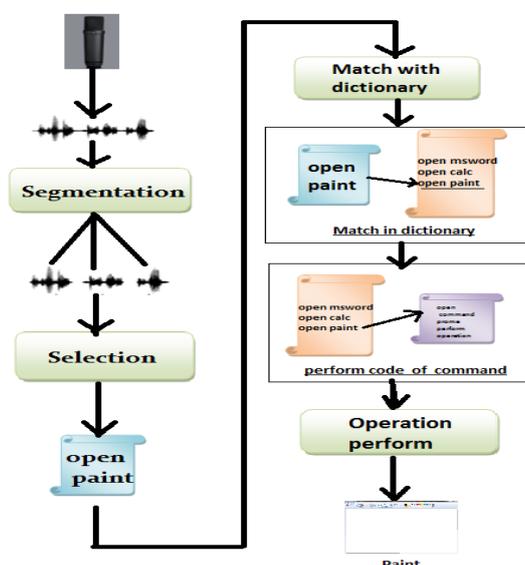
Brief Idea about Voice Recognition:

Fig.3 Brief Idea about Voice Recognition.

Example:

The user gives the exit command and the exit operation is performed.

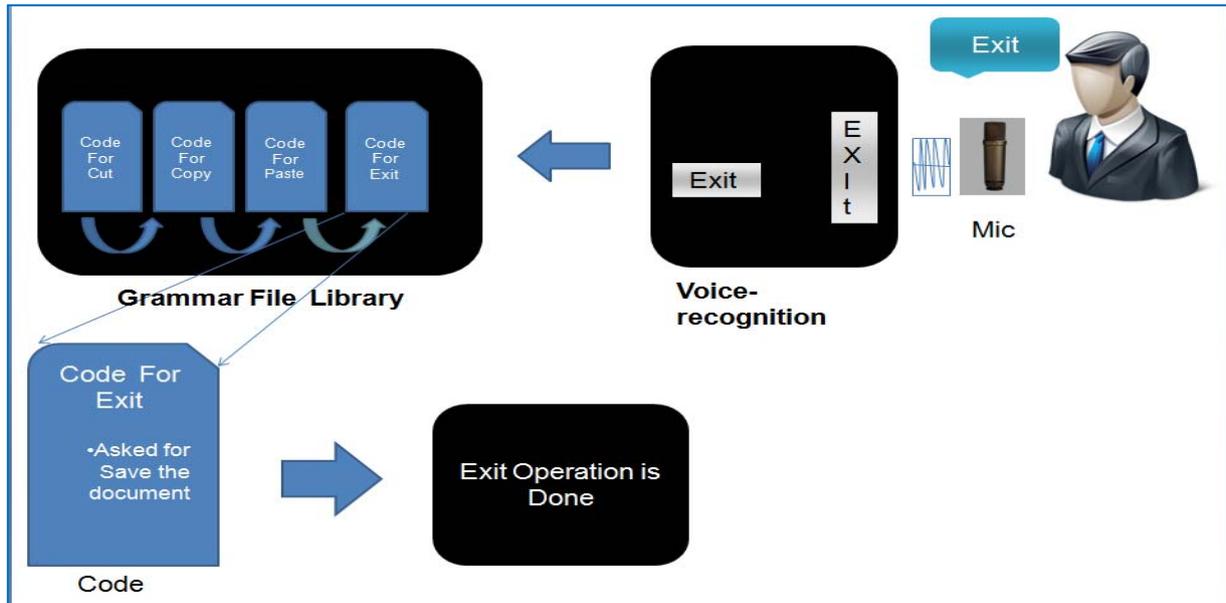


Fig.4 Example

II. RELATED WORK

Existing methodology

1. Analysis of the How Was Your Day?
2. Video lectures converting in text.
3. Indexing for the files.

1) Analysis of the How Was Your Day? [3]

This methodology is based on basic model of voice recognition system. The user give the voice command to system in specific time period .The model is generated each voice samples .The level of voice is analysis and the output that particular model is created like user is sad, user is happy, user in stress. All output of model in day is divided in different class .Analysis all classes and give the overall one result of day.

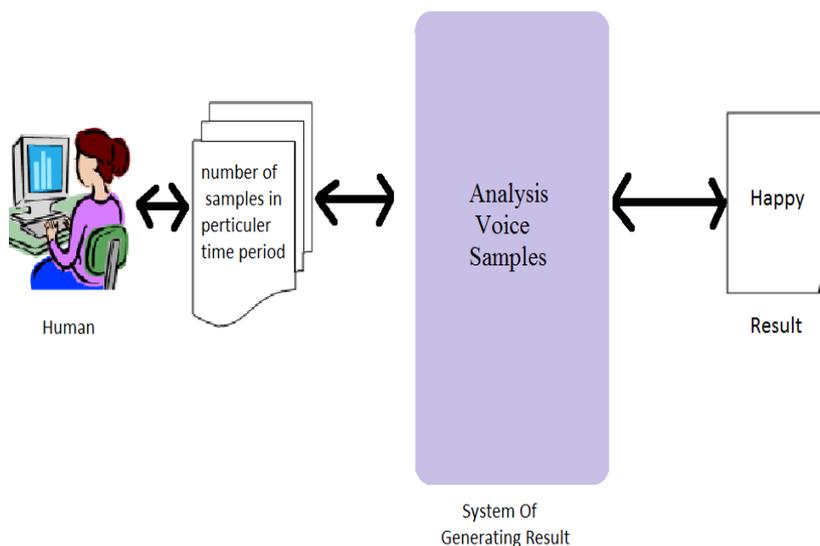


Fig.5 Block diagram of analysis of the How Was Your Day?

The expected output of this methodology is the overall result is display.

2) Video lectures converting in text. [2][1]

In this methodology the video lectures are converted in to text /notes format. The video lectures are converted into audio format and then used voice recognition system to transfer audio to text document.

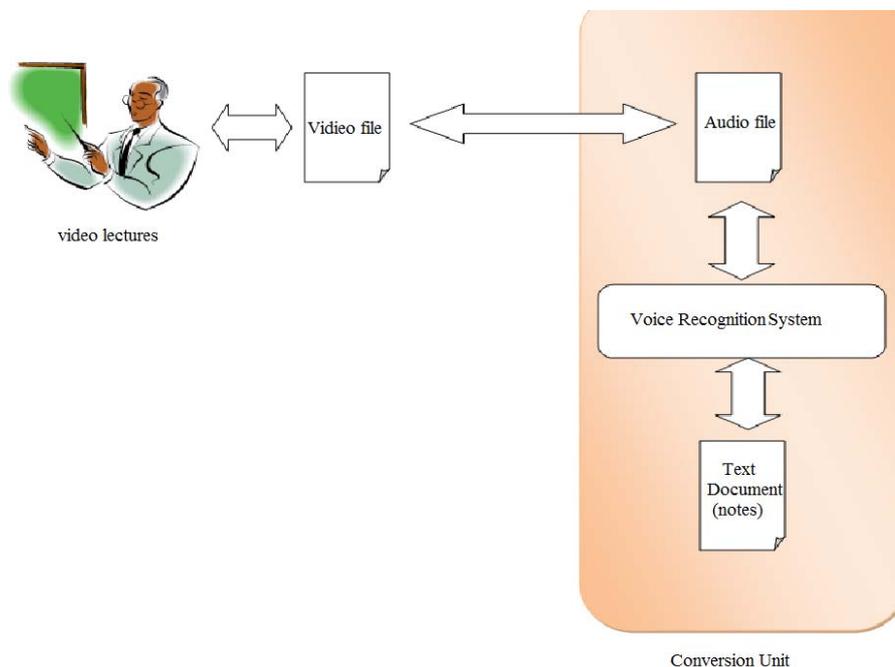


Fig.7 Block diagram of video lectures converting in text

The expected output of this methodology is

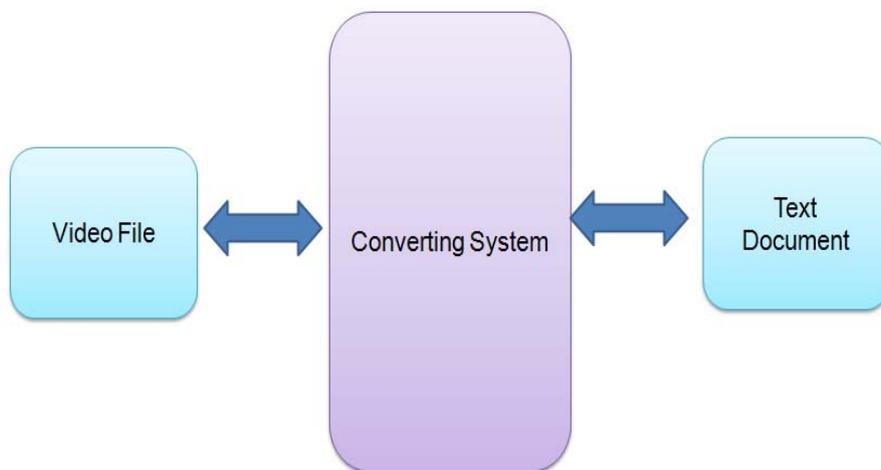


Fig.8 Expected output of video lectures converting in text

3) Indexing for the files [5] [4]

In this model the advanced index feature is added in existing video lectures converting in text methodology. Index is used for the easy searching the video lectures ,audio lectures, notes ,PPT of that lecture .The database is prepare for the lectures .the user is give command and the operation is perform

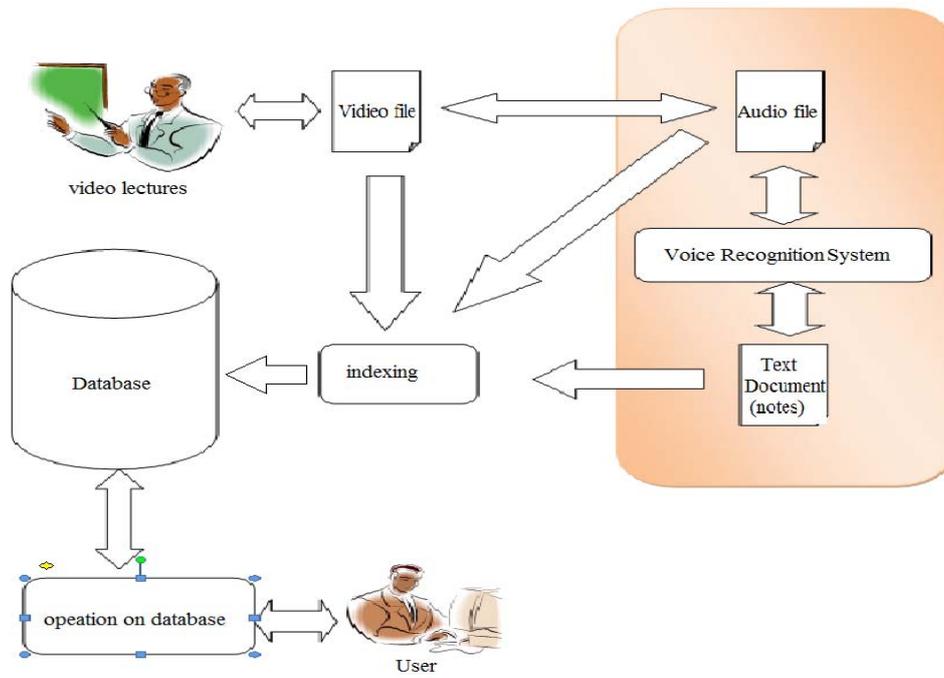


Fig.9 Block diagram of Indexing for the files

The expected output of this methodology is

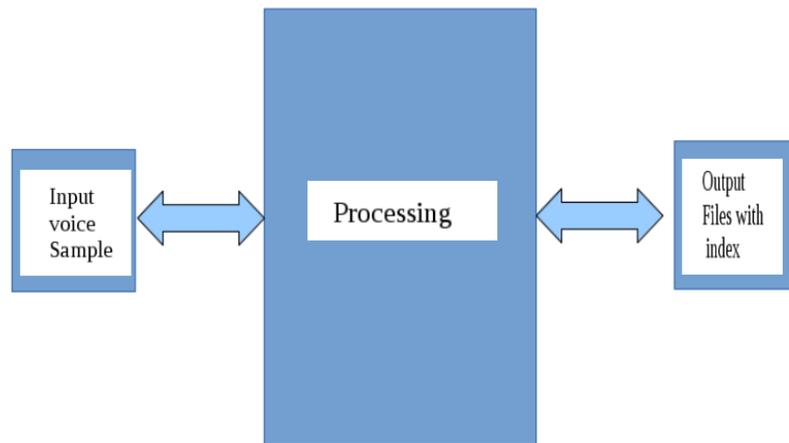


Fig.10 Expected output of Indexing for the files.

III. STATE DIAGRAM FOR COMMAND PROCESS

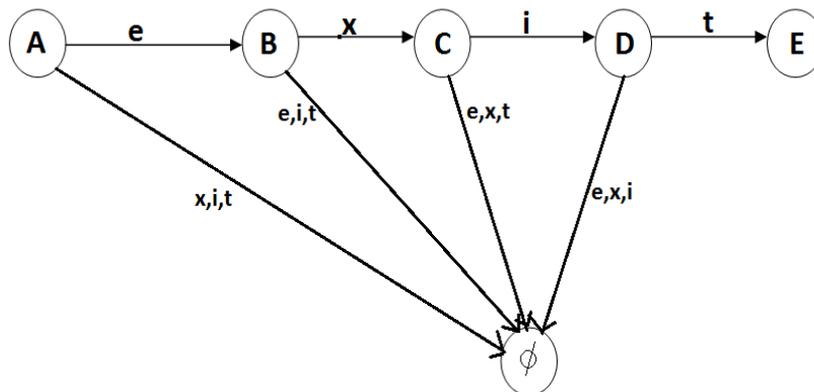


Fig.11.State diagram.

IV. MATHEMATICAL MODEL

Voice recognition: {start, end, input, output, NDD, DD}

Start: Starting state.

Input: {i1, i2, i3, i4....} set of voice characters

Output: command is executed

NDD (non-deterministic data) : {NULL, w1} null or wrong input

DD (Deterministic data) :{d1} proper command is given.

V. CONCLUSION

Analysis of the How Was Your Day? This Existing methodology the basic model of voice recognition is support. This methodology is only limited used .only we can give the voice file the analysis model give the feedback like day is good, bad, happy, sad etc. This methodology is not reusable.

In the second methodology added the advanced voice recognition system which can be reusable and easy to used .This methodology not a limited scope but in this methodology the large data is not supported .large data created a confusion .and the search is not simple in the data. To overcome this disadvantage of this methodology we adding the index to the documents for easily accessing the data .we get appropriate data in minimum time. We perform the add, delete, search operation on data.

References

1. Haojin Yang and Christoph Meinel, Member, IEEE, "Content Based Lecture Video Retrieval Using Speech and Video Text Information", IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES, VOL. 7, NO. 2, APRIL-JUNE 2014 1939-1382 2014 IEEE.
2. Rohit Ranchal, Member, IEEE, Teresa Taber-Doughty, Yiren Guo, Keit Bain, Heather Martin, J. Paul Robinson, Member, IEEE, and Bradley S. Duerstock, "Using Speech Recognition for Real-Time Captioning and Lecture Transcription in the Classroom", IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES, VOL. 6, NO. 4, OCTOBER-DECEMBER 2013.
3. David Benyon, Björn Gamba, Preben Hansen, Oli Mival, and Nick Webb, "How Was Your Day? Evaluating a Conversational Companion", IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, VOL. 4, NO 3, JULY-SEPTEMBER 2013.
4. Vijaya Kumar Kamabathula, Sridhar Iyer, "AUTOMATED TAGGING TO ENABLE FINE-GRAINED BROWSING OF LECTURE VIDEOS".
5. Xuedong Huang and Li Deng, "An Overview of Modern Speech Recognition".