



# **iJRASET**

International Journal For Research in  
Applied Science and Engineering Technology



---

# **INTERNATIONAL JOURNAL FOR RESEARCH**

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume: 6      Issue: IV      Month of publication: April 2018**

**DOI: <http://doi.org/10.22214/ijraset.2018.4606>**

**[www.ijraset.com](http://www.ijraset.com)**

**Call: ☎ 08813907089**

**E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)**

# Automatic Speech Recognition: - A Literature Survey

Nakshatra Sain<sup>1</sup>, Sagar Kapse<sup>2</sup>, Malika Bindal<sup>3</sup>, Surbhi<sup>4</sup>, Prof. Prashant S. Chavan<sup>5</sup>

<sup>1, 2, 3, 4, 5</sup> Department of information Technology, Bharati Vidyapeeth (to be Deemed University) College of Engineering Pune

**Abstract:** Automatic Speech recognition is a technology that enable a computer to capture the words spoken by a human with a help of microphone. These words are later recognized by speech recognizer, and in the end, system outputs the recognized words. The output of the system depends on a lot of things as noisy environment, utterance of word and dialect of the person. It's an effective method which provides ease of use. The technology is widely used these days in smart homes where lightning is controlled with the help of voice. It is also being implemented in car audio system, every mobile phone and high end military aircraft.

**Keywords:** Automatic Speech Recognition, ASR, Siri, Artificial Intelligence, MFCC

## I. INTRODUCTION

### A. Discourse Acknowledgment Framework Comprises of Fundamental Things

Highlight extraction, Acoustic demonstrating, Articulation displaying, Decoder. The procedure of discourse acknowledgment starts with a speaker talking anything before receiver. The waves are gathered and changed over into electrical signs. These electrical signs are then changed over into computerized shape to make them reasonable by the discourse framework. Discourse flag is then changed over into list of highlight vectors, which is accepted to contain just the significant data about given articulation. An imperative property of highlight extraction is the concealment of data unessential for revise arrangement, for example, data about speaker and data about transmission channel (normal for a receiver). At long last acknowledgment part finds the best match in the learning base, for the approaching element vectors. Here and there the future extraction strategy may set aside a long opportunity to stack because of trouble. Highlight extraction techniques like Mel recurrence cepstral coefficient (MFCC) gives some approach to get uncorrelated vectors.

## II. LITERATURE SURVEY

- A. Automated speech recognition is an emerging technique that helps in recognizing the human speech by the machine. Everyday research is going on in building a model for recognizing speech and converting into text. The paper summarizes the various approaches for speech recognition, acoustic phonetic approach, pattern recognition approach and have also explained there benefits and disadvantages. The author has also provided detail work related to the conclusions.
- B. In this, the authors have enhanced the working of system. They have used MFCC and Distance Minimum techniques in the identification phase. These two techniques provided more efficient speaker identification system. The speech recognition phase uses the most efficient HMM Algorithm. Speaker recognition module improves the efficiency of speech recognition scores. The coding of all the techniques mentioned above has been done using MATLAB. They have come to a conclusion that the combination of MFCC and Distance Minimum algorithm gives the best performance and also accurate results in most of the cases with the efficiency of 95%. And that the HMM algorithm is able to identify the balanced isolated word. Speech recognition system achieves 98% efficiency. Autocorrelation technique is best suited for noisy environments. It also takes less memory space and time for computation
- C. Speech acknowledgment is a testing issue to manage. The creators has clarified of how much this innovation has advanced in the earlier years. They have explained how earlier 'Acoustic Phonetic Approach' was used and it was replaced soon by 'Pattern Recognition' in early 70's and how in today's word 'Artificial Intelligence' has enhanced and improved these methods of speech recognition after 90's. Speech recognition is one of the most interesting areas that researchers like to research on as they provide ease of work, the applications of speech recognition is definitely going to increase in upcoming years.

## III. PROPOSED SYSTEM

The Artificial Intelligence approach is a mixture of the acoustic phonetic approach and pattern recognition approach.

Automatic speech recognition system involves two phases: Training phase and recognition phase. A basic training procedure is followed to map the basic speech unit such as phone, syllable to the acoustic observation. In training phase, known speech is recorded, pre-processed and then enters the first stage i.e. Feature extraction.

The next three stages are HMM creation, HMM training and HMM storage. The recognition phase starts with the analysis of unknown speech signal. The signal captured is converted to a series of acoustic feature vectors. The speech is compared against the HMM's networks using appropriate algorithm and the word which is pronounced is displayed. An ASR system can only recognize what it has learned during the training process. But, the system is able to recognize even those words, which are not present in the training corpus and for which sub-word units of the new word are known to the system and the new word exists in the system dictionary.

The speech recognition architecture is built with the help of various major components such as acoustic front end, acoustic model, lexicon, language model and decoder. While recording phase the end user will read all the words which are there in the sentences of the current sets of that application. Then recording will be saved in the database and then later the sentences which we have recorded from the user are compared with stored sentences which were there in the database. And best match is selected. This is called as pattern recognition. There are four basic components are as follows:-

The speech recognition architecture is built with the help of various major components such as acoustic front end, acoustic model, lexicon, language model and decoder. While recording phase the end user will read all the words which are there in the sentences of the current sets of that application. Then recording will be saved in the database and then later the sentences which we have recorded from the user are compared with stored sentences which were there in the database. And best match is selected. This is called as pattern recognition. There are four basic components are as follows:-

#### *A. Acoustic front end*

Acoustic front end is nothing but the signal processing and it also extract features. The main goal of feature extraction step is that to compute a sequence of feature vectors providing with a compact representation of the given input signal. It also takes care about of converting the speech signal into its relevant features which will provides us useful information for recognition.

#### *B. Acoustic Model*

In speech recognition architecture acoustic model is the most important knowledge source for automatic speech recognition which provides us the features for phonetic units to be recognized.

#### *C. Language Model*

It is the collection of constraints on which the sequence of the word is accepted in the form of given language.

#### *D. Decoder*

The word decoder is nothing but the task to find out the most likely word in the given sequence W which was given in the observation sequence O and also in the acoustic phonetic language model. There were various problems while decoding which can be get solved through by using dynamic programming algorithms'.

### IV.CONCLUSIONS

Discourse acknowledgment is a standout amongst the most coordinating regions of machine insight, since people complete a day by day action of discourse acknowledgment. It has pulled in researchers as a vital teach and has made a mechanical effect on society and, is relied upon to prosper facilitate in zone of human machine cooperation. Now with the help of Artificial intelligence the phenomenon has become sorted and with the use of few techniques now the efficiency has become up to 95%, we have focussed on providing the latest trend in speech recognition. Speech Recognition technology definitely has a huge prospects in terms of disabled person and the future may see some great enhancement in the technology.

### REFERENCES

- [1] Vijayalakshmi A, Midhun Jimmy, Moksha Nair, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 Issue 3, March 2015.
- [2] Suma Shankaranand , Manasa , Mani Sharma , Nithya A.S , Roopa K.S. , K.V. Ramakrishnan6International Journal of Recent Development in Engineering and Technology Website: [www.ijrdet.com](http://www.ijrdet.com) (ISSN 2347-6435(Online) Volume 2, Issue 3, March 2014
- [3] Shikha Gupta ,Mr.Amit Pathak ,Mr.Achal Saraf International Journal of Science, Engineering and Technology Research (IJSETR), Volume 3, Issue 8, August 201
- [4] Kuldeep Kumar, Ankita Jain and R.K. Aggarwal, "A Hindi speech recognition system for connected words using HTK", Int. J. Computational Systems Engineering, vol. 1, no. 1, pp. 25-32, 2012.





10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)