IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

# Different Language Recognition Model

**Sankirna D. Joge[1], Prof. A. S. Shirsat[2]**

M.E Student, Dept. of ECE, Smt. Kashibai Nawale College of Engineering, Vadgaon (Bk.), Pune, India[1]

Professor, Dept. of ECE, Smt. Kashibai Nawale College of Engineering, Vadgaon (Bk.), Pune, India[2]

**Abstract:** Language is the ability to know any complex era in a real world application. Approximate number of languages is 6500. It is possible to record the language in either analog or digital media for example, in graphic writing, braille, or whistling because human language is modality-independent. Different regions in a world have different languages spoken. When a human meets another human, speaking different language, it is difficult to identify the one what next person is talking about or in which language. Hence, the main focus is on to detect the language which is spoken by a next person. For this the database of different languages is recorded. Pc as a hardware with Matlab software is used for extracting feature of speech to detect the particular language. For detection of language, database of different languages is created. The signal processing is done on a database so as to get the final output which is final detection of a language.

**Keywords:** Language, Hardware, Recognition, Database, Signal processing, etc.

## I. INTRODUCTION

Speech is very popular for interacting between man and machines. In signal processing, speech recognition can be done according to application. It is a verbal way of communication and this research area is full of innovative approaches [6]. There are many applications for speech recognition like virtual gaming, security, sign language interpretation. Human speech is read by an input sensing device like mike [1]. The human speech is needed to recognize in the speech processing.

Using a speech recognition system, it is possible to design a system that can detect a language that has been spoken. Since 1990s; people have moved their interests to the difficult task of Large VocabularyContinuous Speech Recognition (LVCSR) and indeed achieved a great progress. Speechrecognition has been developed from theoretical methods to practical systems [9]. Now a days Google is providing Voice Search facility forweb searching and android mobile is providing one more facility i.e. Speech-to-Textconversion for messaging [4].

## II. CHALLENGES

The aim of this project is to efficiently design a system in Matlab that will detect any language spoken by a person by using efficient speech recognition techniques and also to improve its recognition accuracy.

The basic objective of a project is to implement a system which can recognize multiple languages. For this, the following objectives are:

- To design and implement a system on Matlab with PC/Laptop as a hardware platform.
- To collect a database for training of a system.
- To select proper classifier for classification.
- To detect at most three languages with good accuracy.

## III. METHODOLOGY

Block Diagram

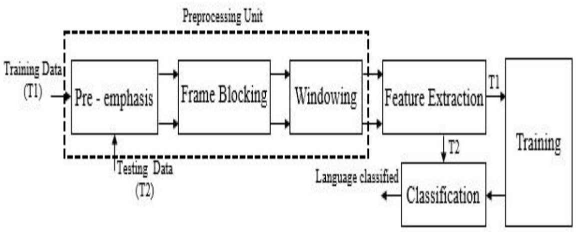


Fig. 1: Speech Recognition and its Interpretation

Fig.1 shows block diagram of the speech processing for language detection. First speech is taken by an input sensing device like mike then acquisition of that speech signal is taken place. Endpoint Detection (EPD) algorithm is used in speech pre-processing part to eliminate background noise, after that feature have to be extracted from it and Mel-frequency Cepstral coefficient (MFCC) algorithm is used for this feature extraction purpose. For language detection, i.e. for classification, multi-class SVM is used [1].

### A. Mike

A microphone or mike is an acoustic-top-electric transducer. It converts the speech signal into an electrical signal. It has in-built data acquisition system and amplifier which is necessary to amplify signal [1].

Fig.2: Electrical Microphone [4].

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

### B. Speech Acquisition

Speech (data) acquisition comes under feature extraction analysis. The challenge is about to handle the environmental noise including interfering speakers (background noise) and convolution distortion (noise due to less expensive microphone and data acquisition) [9]. The input speech is recorded in MATLAB environment in 1sec of duration with sampling rate of 8 kHz. Voice frequency band ranges from approximately 300 Hz to 3400 Hz. Suppose the highest frequency component, in hertz, for a given analog signal is $f_{max}$ .According to the Nyquist theorem, sampling rate must be at least $f_{max}$ , or twice the highest of frequency component present in a signal [1].

$$fmax = 3400 \text{ Hz or } 4 \text{ KHz}$$
$$\text{Sampling rate} = 2 \times 4K = 8 \text{ KHz}$$

### C. Pre-Processing

An important pre-processing step in Automatic Speech Recognition systems is to detect the noise [1]. Word recognition accuracy is improved by the end point detection algorithm. If the region of the speech is well known into a signal, then processing time get reduced. This aspect is also very important for recognition. Thus, for developing speech recognition device capable of working almost in all environments an appropriate endpoint detection algorithm is needed [8].

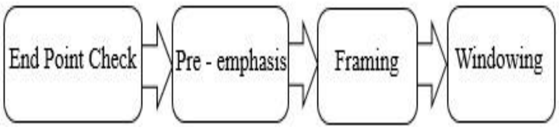

Fig.3: Pre-processing construction [8].

Endpoint checking can find the head and tail of useful signal, pre-emphasis can reduce the signal dynamic range, framing can divide the speech data into small chunks. The window is added in order to improve the speech processing. The endpoint detection problem is nontrivial for non-stationary backgroundnoise where artefacts (i.e., on speech events) may be introduced by the speaker and the recording environment [5].

### D. Feature Extraction

In order to represent the static acoustic properties, the Mel-Frequency Cepstral Coefficient (MFCC) [3] is used as the acoustic feature in the cepstral domain. This is a fundamental concept which uses a set of non-linear filters to approximate the behavior of the auditory system. It adopts the characteristic of human ears that human is assumed to hear only frequencies lying on the range between 300Hz to 3400Hz. Human's ears have high resolution to the low frequency. Those are more sensitive. So, the filter bank is designed to emphasize the low frequency over the high frequency. The Mel-Frequency Cepstral Coefficients (MFCC) is used for feature extraction in Automatic Speech Recognition technique (ASR) [2].

### E. Classification

For the multi-language detection, multi-class SVM is designed. Multi-class SVM is designed from a binary classification. For designing the multi-class SVM "One-vs-all" strategy is used. Benefit of multi-class SVM is interpretability. In "One-vs-all" strategy one classifier per class is an approach [3]. Decision tree, Bayesian tree are mainly used for multi-classification. The learning phase of the classifiers is done using as training data only a subset of instances from the original data-set which contain any of the two corresponding class labels.

## IV. SPECIFICATION

### A. Software

The project is PC based on Matlab 2010b environment which is used to design the speech recognition system. At the end of the project, GUI of a complete system is designed through which we can train and give the input to the program [4].

### B. Hardware

The project is designed on a PC through Matlab software. The software requires minimum 1GB RAM. x86 or x64 based Intel or AMD Processor. Minimum operating system requirement is Windows XP [4].

TABLE I : RESOURCE UTILIZATION IN A PROJECT

| Sr. No. | Resource Name | Resource Used | Utilization |
|---|---|---|---|
| 1 | Matlab (R2010a) Version : 7.10.0.499 | Own Development Environment | - |
| 2 | Processor Family | Intel Core i5 | - |
| 3 | RAM | 238 MB/6 GB | 4 % |
| 4 | Total Connectors | 1/9 | 11.11 % |
| 5 | Headphone with Mic | - | - |

## V. RESULT

The testing of the project is being done with three languages viz. English, Marathi and Hindi. The various samples of these three languages are given to the trainer though live recording or .wav files input.

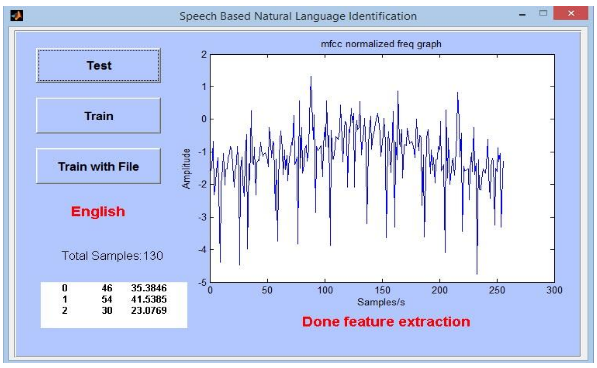

Fig.4: Output Window in Matlab

In a Fig.4 the spoken sentence "Hi, how are you?" is tested. After 3 seconds of input, "English" as an output is detected with MFCC graph.

TABLE III: ACCURACY FOR EACH LANGUAGE SPOKEN

| Language | Total Sentence Spoken | Total Sentence Detected | Accuracy |
|---|---|---|---|
| English | 10 | 9 | 90 % |
| Marathi | 10 | 7 | 70 % |
| Hindi | 10 | 8 | 80 % |
| **Total** | **30** | **24** | **80%** |

The accuracy per word is calculated each time a word spoken. Overall accuracy is 80 % for SVM multi class classifier.

## VI. CONCLUSION

The spoken words for Hindi, Marathi and English language are being successfully detected. The accuracy of the system depends on number of unique and clear samples given. More the number of samples more will be accuracy. The use of deep neural network for classification can significantly increases accuracy. However, the samples required per language should be minimum 500. The SVM Multi-Classifier can classify more than 2 objects successfully with moderate number of samples. The accuracy of the overall project is 80 %.

## ACKNOWLEDGMENT

It is my pleasure to get this opportunity to thank my beloved and respected Guide **Prof. A. S. Shirsat** who imparted valuable basic knowledge of Electronics and telecommunications specifically related to Signal Processing Domain. We are grateful to Elec. &Comm. Smt. Kashibai Nawale College of Engineering, Pune for providing us infrastructure facilities and moral support.

## REFERENCES

[1] Xin Guang Li, Min Feng Yao and JiaNeng Yang, "Speech Recognition Approach Based on Speech Feature Clustering and HMM" Journal of Computers, Vol. 7, No. 9, pp. 2269-2275, September 2012

[2] K.Saksamudre, Suman, P.P. Shrishrimal, and R.R. Deshmukh. "A Review On Different Approaches for Speech Recognition System". International Journal of Computer Applications, Vol. 115, No.22, pp.24-28, April 2015.

[3] Michelle Cutajar, Edward Gatt, Ivan Grech, Owen Casha, Joseph Micallef, "Comparative study of automatic speech recognition techniques", IET Signal Process, Vol. 7, Issue. 1, pp. 25–46, August 2013

[4] M. Lutter, "The Speech Recognition Wiki", Recognize-speech.com,2014.[Online]. Available:http://recognize-speech.com/feature-extraction/mfcc. [Accessed:15-Nov- 2015].

[5] Kashyap Patel and R.K.Prasad "Speech Recognition and Verification using MFCC & VQ", International Journal of Emerging Science and Engineering, vol. 1, no. 7, pp. 33-37, May 2013.

[6] C. Ittichaichareon, S. Suksri and T. Yingthawornsuk, "Speech Recognition using MFCC", International Conference on Computer Graphics, Simulation and Modeling, vol. 2, no. 8, p. 4, July 2012.

[7] A. Bala, A. Kumar and N. Birla, "Voice Command Recognition System Based On MFCC and DTW", International Journal of Engineering Science and Technology, vol. 2, no. 12, pp.7335-7342, March 2010.

[8] S. Deshmukh and M. Bachute, "Automatic Speech and Speaker Recognition by MFCC, HMM and Vector Quantization", International Journal of Engineering and Innovative Technology(IJEIT), vol. 3, no. 1, p. 6, July 2013.

[9] V. Tiwari, "MFCC and its Applications in Speaker Recognition", International Journal on Emerging Technologies, vol. 1, no. 1, pp.19-22, February 2010.