*International Journal of Engineering & Technology*, 7 (2.33) (2018) 426-428

International Journal of Engineering & Technology

*Website: www.sciencepubco.com/index.php/IJET*

***Research paper***

# Simulation and detection of tamil speech accent using modified mel frequency cepstral coefficient algorithm

**Swagata Sarkar [1] *, Sanjana R [2], Rajalakshmi S [2], Harini T J [2]**

*[1] Assistant Professor, Sri Sairam Engineering College*
*[2] Student, Sri Sairam Engineering College*
**Corresponding author E-mail: swagata.ei@sairam.edu.in*

**Abstract**

Automatic Speech reconstruction system is a topic of interest of many researchers. Since many online courses are come into the picture, so recent researchers are concentrating on speech accent recognition. Many works have been done in this field. In this paper speech accent recognition of Tamil speech from different zones of Tamilnadu is addressed. Hidden Markov Model (HMM) and Viterbi algorithms are very popularly used algorithms. Researchers have worked with Mel Frequency Cepstral Coefficients (MFCC) to identify speech as well as speech accent. In this paper speech accent features are identified by modified MFCC algorithm. The classification of features is done by back propagation algorithm.

***Keywords***: *Artificial Neural Network; Back Propagation Network; Mel Frequency Cepstral Coefficients; Speech Accent.*

## 1. Introduction

Speech processing is a very important topic used for speaker recognition as well as speech accent recognition. In literature review different algorithms and methods are discussed. Robert Rehr, Timo Gerkmann [1] enhance the speech using Machine Learning Spectral Envelopes (MLSE) based approaches with Super Gaussian Speech priors. Narendra D. Londhe, Ghanahshyam B. Kshirasagar [2] characterize speech using Mel frequency Cepstral Coefficients (MFCC) and Cross Validation Technique to improve the performance. Anand H. Unnibhavi, D.S.
Jangamshetti [3] extracts the feature by using Linear Predictive Coding (LPC) and classified using Euclidean Distance.Sanjay Bhardwaj, Sunil Pathania, Rajesh Akela [4]
uses Hidden Markov Model (HMM) for classification and for best Partial path identification Viterbi Algorithm and Euclidean Distance are used. Kamil Kaminski, Ewelina Majda, Andrzej P. Dombrowski [5] uses Voice Print (VP) technique to describe voice and Gaussian Mixtures Model (GMM) for classification process. Patil, S.D. Shirbahadurkar, A.N. Paithane [6] uses Mel Frequency Cepstral Coefficients (MFCC) for feature extraction and Vector Quantization (VQ) and Gaussian Mixtures Model (GMM) for classification.
Feature extraction techniques such as LPC, MFCC are used. In this way training vector is generated from speech signal of utterance spoken by the user. These test patterns are tested against training pattern using different pattern classification technique such as HMM, GMM, Viterbi Algorithm, VQ and ANN.

## 2. Proposed model

MFCC is used for both speech Recognition and accent detection. In this paper, Modified MFCC is used for feature extraction.

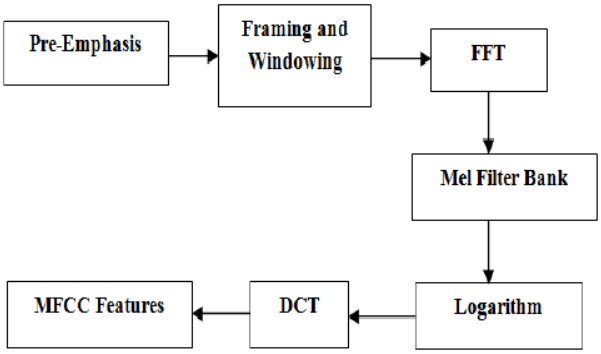


**Fig. 1:** MFCC Block Diagram.

Pre-emphasis is performed for flattening the magnitude spectrum and balancing the high and low frequency components. Windowing function weights the signal in the time domain and divides it into a sequence of partial signals. There are many windowing techniques such as Hamming Window, Hanning Window, Rectangular Window, Triangular window, Blackman Window. For better efficiency, in this paper Windowing Technique is modified. Instead of using default hamming window in MFCC algorithm, Hanning window is used for feature extraction. The Fast Fourier Transform (FFT) is a basic technique for digital signal processing applicable for spectrum analysis. FFT is often used to compute numerical approximations to continuous Fourier. Discrete Cosine Transform (DCT) is applied to the filter banks retaining a number of the resulting coefficients while the rest are discarded. MFCC mimics the logarithmic perception of loudness and pitch of human auditory system and tries to eliminate speaker dependent characteristics by excluding the fundamental frequency and their harmonics. To represent the dynamic nature of speech the MFCC also includes the change of the feature vector over time as part of the feature vector. The extracted features are classified with back propagation algorithm of Artificial Neural Network.

## 3. Results and discussions

In this paper Modified MFCC algorithm is used for speech feature extraction. In this algorithm speech features are extracted based on Hanning windowing techniques. In figure 2 Hanning window is shown as an example.

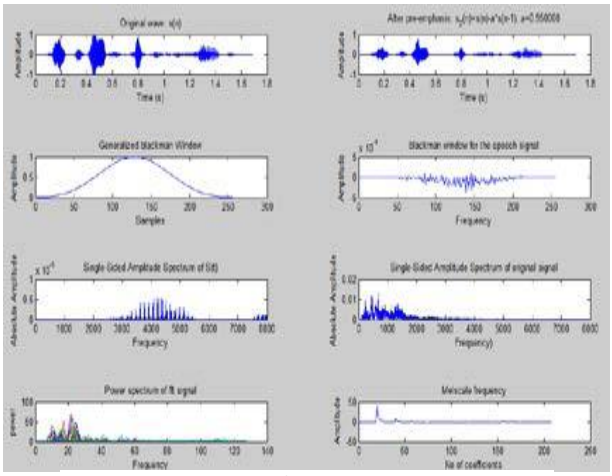

**Fig. 2:** Hanning Window for Speech Accent Recognition.

Thirteen coefficients are identified as feature set at the output of the modified MFCC. Thirteen features are used as the input of artificial neural network and four outputs are taken from the output layer. The architecture of the ANN is 13 input neurons, four hidden neurons and four output neurons. Threshold value is considered as 0.01. Sigmoid activation function is used. The example dataset is shown in table 1. The following table represents 5x5 coefficients out of 13 x 13 coefficients of Chennai Tamil accent.
Table 2 shows the data set of Madurai Tamil accent.
The coefficients are processed and normalized to use as the input of artificial neural network. The classification outputs are shown in table 3.
The above results are got for 11320 iterations. The example error versus number of iterations graph are shown in figure 3.

**Table 1:** Chennai Tamil Accent coefficients

| Serial No. | Data for Coefficients | | | | |
|---|---|---|---|---|---|
| 1 | 1.35E-10 | 7.18E-10 | 5.64E-09 | 9.21E-09 | 9.22E-09 |
| 2 | 1.26E-10 | 5.41E-10 | 1.85E-09 | 1.14E-09 | 1.45E-09 |
| 3 | 1.91E-10 | 4.91E-10 | 1.76E-09 | 1.74E-09 | 8.91E-10 |
| 4 | 4.01E-10 | 7.04E-10 | 4.05E-09 | 1.93E-09 | 2.32E-09 |
| 5 | 5.81E-10 | 6.00E-10 | 5.64E-10 | 2.93E-10 | 2.62E-09 |

**Table 2:** Madurai Tamil Accent Coefficients

| Serial No. | Data for Coefficients | | | | |
|---|---|---|---|---|---|
| 1 | 7.07E-08 | 1.11E-10 | 4.04E-08 | 5.43E-10 | 0 |
| 2 | 6.85E-08 | 7.88E-08 | 2.69E-08 | 5.43E-10 | 0 |
| 3 | 2.26E-08 | 1.64E-08 | 1.88E-08 | 5.43E-10 | 0 |
| 4 | 1.70E-07 | 8.00E-08 | 5.57E-08 | 5.42E-10 | 0 |
| 5 | 3.48E-07 | 4.66E-08 | 6.76E-08 | 5.42E-10 | 0 |

**Table 3:** Classification Outputs

| Serial No. | Kovai | | Chennai | | Tirunelveli | | Madurai | |
|---|---|---|---|---|---|---|---|---|
| | Expected Output | Actual Output | Expected Output | Actual Output | Expected Output | Actual Output | Expected Output | Actual Output |
| 1 | 1 | 0.9872 | 0 | 0.0117 | 0 | 0.0072 | 1 | 0.9923 |
| 2 | 0 | 0.0042 | 0 | 0.0200 | 1 | 0.9820 | 1 | 0.9974 |
| 3 | 0 | 0.0127 | 1 | 0.9885 | 1 | 0.9925 | 0 | 0.0079 |
| 4 | 0 | 0.0115 | 0 | 0.0008 | 0 | 0.0125 | 1 | 0.9874 |

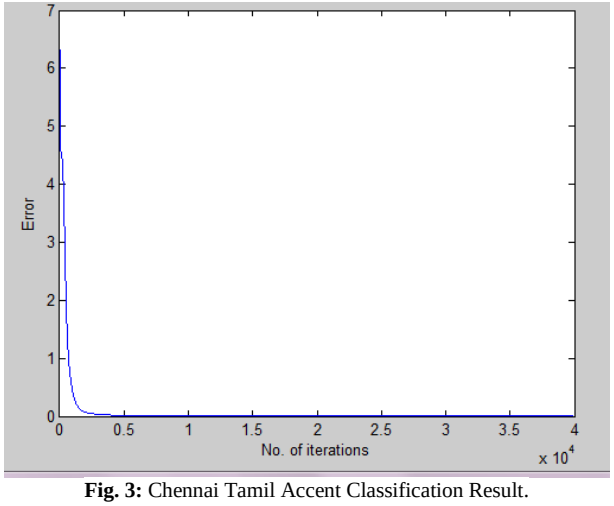


**Fig. 3:** Chennai Tamil Accent Classification Result.

## 4. Conclusion

In this paper modified MFCC algorithm is used instead of MFCC algorithm for feature extraction. Further features are classified using Back Propagation Network. It is seen that each language accent is identified efficiently with minimum error.

## Acknowledgement

Authors are greatly privileged to get a wonderful working environment in the institution.

## References

[1] Robert Rehr, Timo Gerkamnn (2018), "On the Importance of Super Gaussian Speech Priors for Machine Learning Based Speech Enhancement", *IEEE*.

[2] Narendra D. Londhe, Ghanahshyam B. Kshirasagar (2017), "Speaker Independent Isolated Words Recognition System for Chhattisgarhi Dialect", *ICIIECS*.

[3] Anand H. Unnibhavi, D.S. Jangamshetti (2017), LPC Based Speech Recognition for Kannada Vowels, *ICEECCOT*.

[4] Sanjay Bhardwaj, Sunil Pathania, Rajesh Akela (2015), Speech Recognition using Hidden Markov Model and Viterbi Algorithm, *IJARECE*.

[5] Kamil Kaminski, Ewelina Majda, Andrzej P. Dobrowolski (2013), Automatic Speaker Recognition using a unique personal feature vector and Gaussian Mixture Models, *IEEE*.

[6] U.G. Patil, S.D. Shirbahadurkar, A.N. Paithane (2016). Automatic Speech Recognition of Isolated Words in Hindi Language using MFCC, *CAST*.