International Journal of Research

International Journal of Research (IJR)

e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 11, November 2015
Available at http://internationaljournalofresearch.org

Music Genre Classification using Amplitude and frequency Variants of MFCC

Ashish Sharma* & Mr. Abhishek Tomar**

*Research Scholar, Baddi University Solan, Himachal Pradesh.

**Assistant Professor CSE Deptt, Baddi University Solan, Himachal Pradesh.

ABSTRACT

Music is a good domain for computational recognition of auditory events because multiple instruments are usually played simultaneously. The difficulty in handling music resides in the fact that signals (events to be recognized) and noises (events to be ignored) are not uniquely defined. This is the main difference from studies of speech recognition under noisy environments. Musical instrument recognition is also important from an industrial standpoint. The recent development of digital audio and network technologies has enabled us to handle a tremendous number of musical pieces and therefore efficient music information retrieval (MIR) is required. Musical instrument recognition will serve as one of the key technologies for sophisticated MIR because the types of instruments played characterize musical pieces; some musical forms, in fact, are based on instruments, for example "piano sonata" and "string quartet." Despite the importance of musical instrument recognition, studies have until recently mainly dealt with monophonic sounds. Although the number of studies dealing with polyphonic music has been increasing, their techniques have not yet achieved a sufficient level to be applied to MIR or other real applications. This paper investigates musical instrument recognition using two main approaches like MFCC and BPNN. BPNN have good learning rate in comparison to other classifiers. The whole experiments have been taken place in MATLAB and achieved accuracy is around, FAR= .017 FRR =.018 and accuracy= 100%.

KEYWORDS: MFCC; Music Recognition; Neural Network

1. INTRODUCTION

In recent years with the increasing of multimedia data on theinternet and multimedia databases, there is strong demand for better procedures for automatic classifying, indexing and retrieving multimedia data [1]. This is particularly true for music files, since more and more online music stores and music search applications are available, and they need a more efficient mechanism to organize these massive data. There are two main approaches in indexing and retrieval music files in large scale database: (a) keyword-based indexing annotated by human-beings and (b) content-based indexing and retrieval by automatic classification and labeling. The latter one is more useful since it can save laborious and time–consuming work [2].

Music genre, which is created and utilized by human for categorizing and describing the style property of music files, is one of the most natural attribute for users to utilize to explore in large music database[1]. A group of researchers have made great effort to seek ways of automatically classifying music genres since last decades [1, 2, 3, 4].

These approaches could be considered to be a typical classification task. However, only a single classifier adopted in each system seems to meet the limit of classifier's ability [5].

For this reason, we consider that rather than trying to develop new classifiers, it is better to take the existing ones and combine them in an appropriate way to yield an overall improvement of performance.

A method to integrate classifier together with MFCC to improve classification accuracy, has been successfully adopted in many areas, such as face recognition, handwriting recognition, and speaker recognition. The reason why fusion can improve performance is quite simple: Each of individual classifier may produce some errors, and these errors are assumed to be uncorrelated [6, 7, 8, 9].

In this paper, we adopt a new feature set, including Mel Frequency Cepstral Coefficients (MFCC) and various featuresfrom .wav audio descriptor. The Genetic Algorithm (GA) is used to perform featurereduction. A group of frames are combined together toForma segment, in which mean and variance values of these shorttime frames are calculated. For classifying, we choose neural network classifier fusion method to gain better results.

The outline of this paper is organized as follows. The related work is discussed in section 2. In section 3, various commonly used methods are introduced. Music classification has been implemented in section 4. Experiment and result are given in section 5. Finally, conclusions are drawn in section 6.

2. RELATED WORK

Author	Title	Technique used
Cory McKay, 2010 [10]	Automatic music classification with jMIR	This thesis focuses on the development of jmis, a suite of influential, flexible, easy to get to and unique software tools that can be used to intend, split and apply a broad range of regular music classification technology.
Han Ju, Jian-Xin Xu, and Antonius M.J. Van Dongen, 2010 [11]	Classification of musical styles using liquid state machines	The extremely prearranged chronological arrangement of music suggests it supposed to be adaptable to analysis by a novel spike neural network paradigm: the liquor state machine.
Aziz Nasridinov1 and Young-Ho Park, 2014[12]	A Study on Music Genre Recognition and Classification Techniques	They present a study on techniques for automatic music genre credit and categorization. We first explain machine knowledge based chord credit methods, such as concealed Markov models, neural networks, lively Bayesian network and rule based method, and stencil matching methods. We then clarify supervise, unverified and semi supervised classification methods classify music genres.
Shaveta Sharma, Parminder Singh, 2015 [13]	Extracting GFCC Features for Emotion Recognition from Audio Speech Signals	This presents the completion of the Gamma tone frequency cepstral coefficients filter base characteristic along with BPNN and the untried results on English speech data.

3. RUDIMENTARYPROPOSED METHODS

3.1 Mel Frequency Cepstral Coefficients (MFCC)

Mel Frequency Cepstral Coefficients (MFCCs) are a feature widely used in automatic speech and reciter acknowledgement. They were presented by Davis and Mermelstein in the 1980's, and also have existed state-of-the-art ever since.

International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 11, November 2015
Available at http://internationaljournalofresearch.org

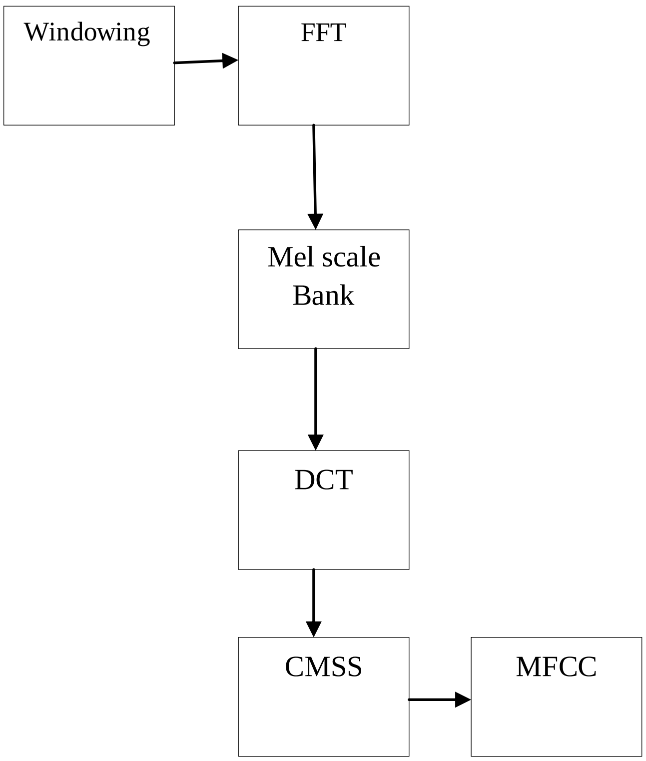


Figure: 1 Block diagram of MFCC

3.2 Genetic Algorithm (GA)

Genetic Algorithms are adaptive heuristic search algorithm based on the evolutionary ideas of normal range and inheritance. As such they signify an intelligent operation of an arbitrary search used to solve optimization problems. Even if randomized, GAs are by no means random, as a substitute they develop past in sequence to direct the search into the region of better act within the search space. The basic techniques of the GAs are calculated to suggest processes in natural systems required for growth; especially those follow the values first laid down by Charles Darwin of "survival of the fitting." Ever since in nature, struggle amongst individuals for scanty income results in the fittest individuals dominate over the weaker ones.

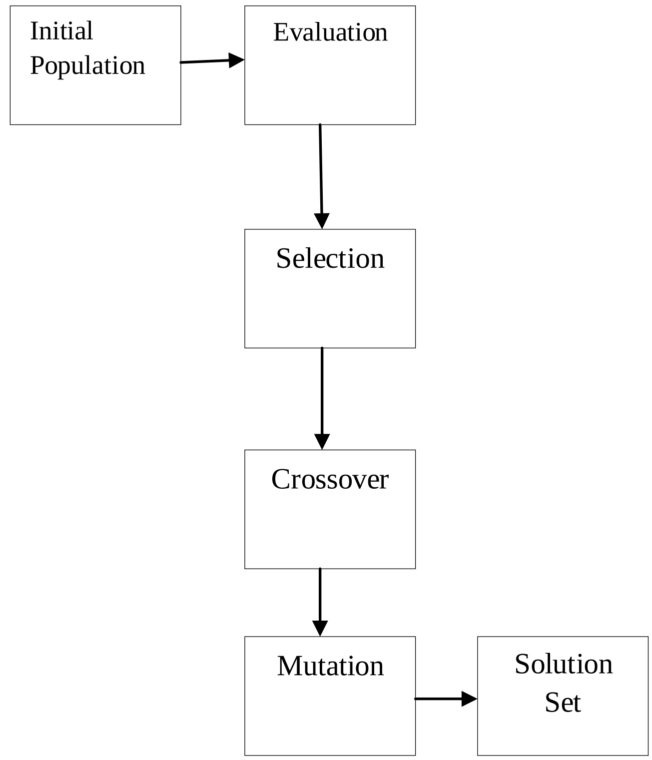


Figure: 2 Block diagram of GA

3.3. Neural Network (NN)

BPNN is used for solving many problems by using the simple output elements. It is the mostly used learning algorithm in the neural network. BPNN is used with fuzzy encoder for understanding the human like reasoning activities of the fuzzy logic system.

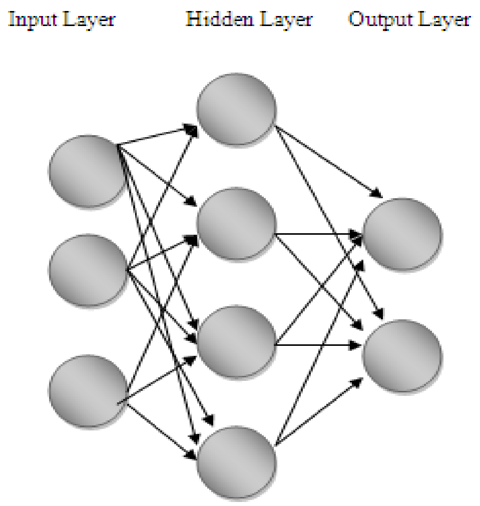


Figure: 3 Back Propagation Neural Network

4 MUSIC GENRE CLASSIFICATION SYSTEM

A. Methodology

4.1 Uploading of audio Files

Audio file of each category will be uploaded for feature extraction. Below algorithm is the process of uploading of audio fileswrite the algorithms in Pseudocodes with expl

```
fullpath=strcat(pathname,'5
(',num2str(i),').wav');

    [x,fs1,N]=wavread(fullpath);

    x=x';

    x=x(1,:);

    x=abs(x);
```

4.2 Training using Neural Network (BPNN)

Below functions will be used by neural network for training of audio features:

```
net=newff(Training_set',Target,20);

net.trainParam.epochs=50;

net=train(net,Training_set',Target);
```

4.3 Feature extraction using MFCC

In most of the Genre categorization MFCC is the nearly all common mark extraction system that is based on the short period frame. In this section, we will use MFCC for feature extraction of the music genre.

4.4 Feature Reduction using GA

Obtained feature set has large number of values, so it becomes very difficult in classification process. So Genetic algorithm will be utilised for the reduction of the feature set. The optimization algorithm for reduction of features can be described below:

```
Function [f] = fitness_fn(e,Fs,Ft)

% Fs= each feature

% Ft = total number of features

% e is classificatiom error rate
(optimization parameter (unknown))

if Fs<Ft

  f=1;

else

   f=0;

end

end

% f=(1-e)*((1-Fs)/Ft);
```

International Journal of Research

International Journal of Research (IJR)

e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 11, November 2015

Available at http://internationaljournalofresearch.org

4.5 Classification using BPNN

After feature extraction process, Classification of music will be done using neural network model. Below algorithm shows the classification process of music signals:

```
testing_row=round((total_rows)*70/100);

testing_set=GareducedFeatures(1:testing_row
);

fori=1:testing_row

target(i)=i;

end

net=newff(testing_set,target,5);

net.trainParam.epochs=50;

net=train(net,testing_set,target);
```

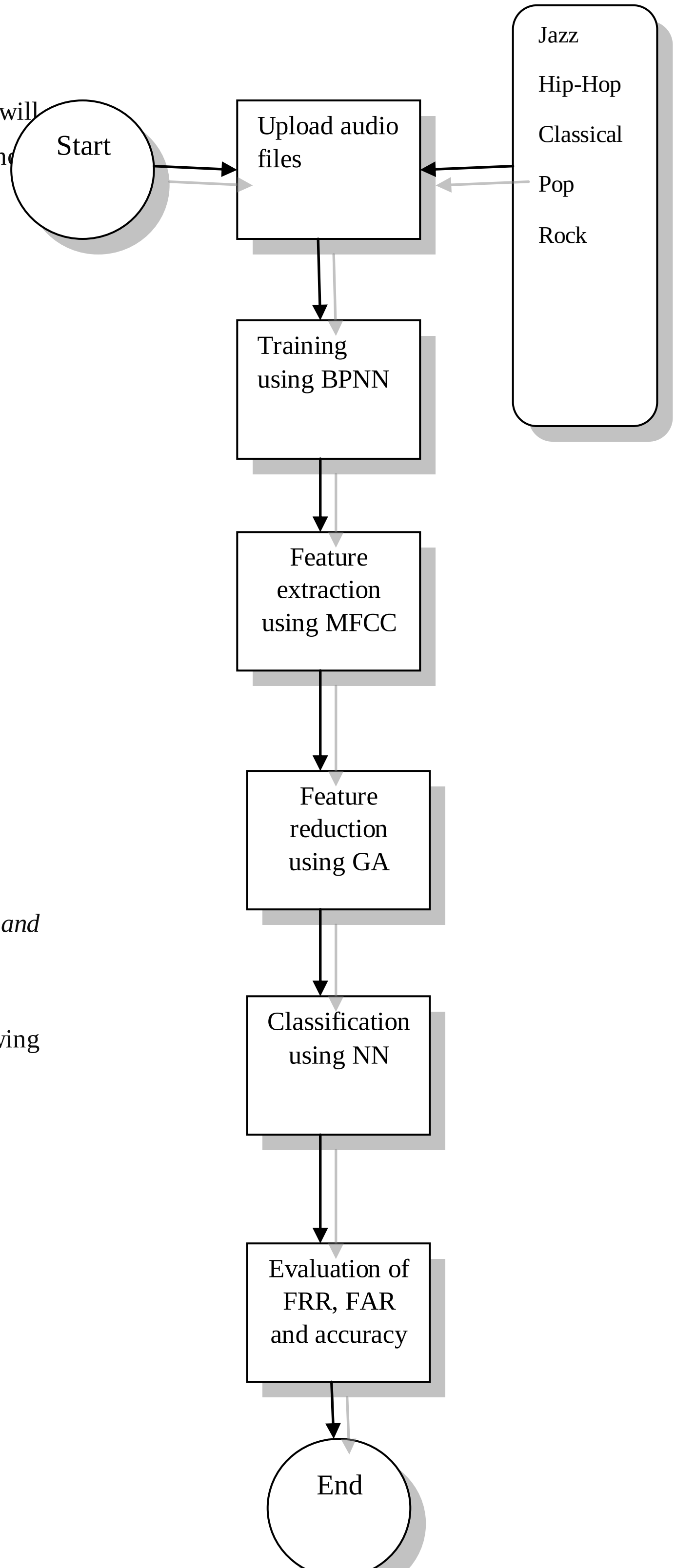
4.6 Result Evaluation on the basis of FAR, FRR and Accuracy

In the end result evaluation will be done using following formulas:

FAR:
far=(error-result)/((numel(test_set)*numel(testing_set))*118);
FRR:frr=(error-far)/((numel(test_set)*numel(testing_set))*118);
Accuracy: accuracy=(1-(far+frr))*100;

B. Flowchart

Below flowchart represents the classification system:



International Journal of Research

International Journal of Research (IJR)

e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 11, November 2015

Available at http://internationaljournalofresearch.org

Figure: 4 Flowchart

5. RESULTS AND ANALYSIS

The whole implementation has been done in MATLAB 2010a environment for the recognition of audio signals. Below graphs and snapshots will show the result analysis of music genre classifications.

Firstly we will upload the music file randomly. Then we set the noise level because we assume that the signal is not noise free signal. Then we will extract the features using MFCC algorithm which is used for feature extraction. It includes Fast Fourier transformation used to convert the time domain signal to frequency domain for spectral analysis, filtration process like hamming window which is a type of filter to attenuate the unwanted frequencies and accepts the required frequency to boost up the frequencies and Error rectangular bandwidth which is the process of bandwidth approximation and to increase the strength of the signal in noisy environment. MFCC also includes filter bank which is an array of Band Pass Filter that separates the input signal into multiple components, each one carrying a single frequency sub-band of the original signal.

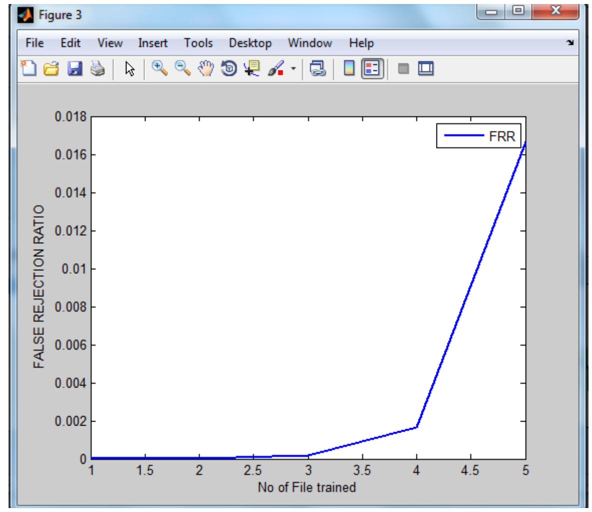


Figure:5FRR Analysis

FRR is the false rejected values. From graph it has been seen that obtained average value is .017.

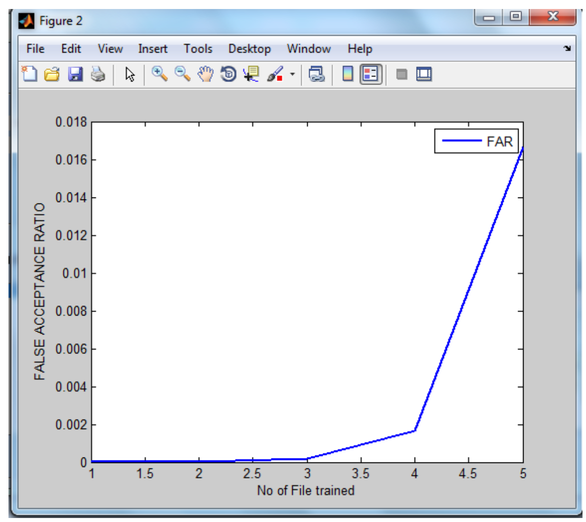


Figure:6FAR Analysis

FAR is the false accepted values. From graph it has been seen that obtained average value is .018.

International Journal of Research
International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 11, November 2015
Available at http://internationaljournalofresearch.org

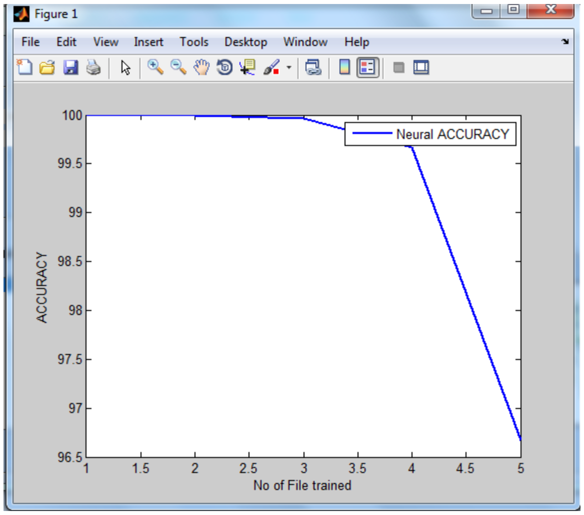


Figure :7Accuracy Analysis

Accuracy is the summation of FAR and FRR. From above figure obtained average accuracy value comes out to be 100%.

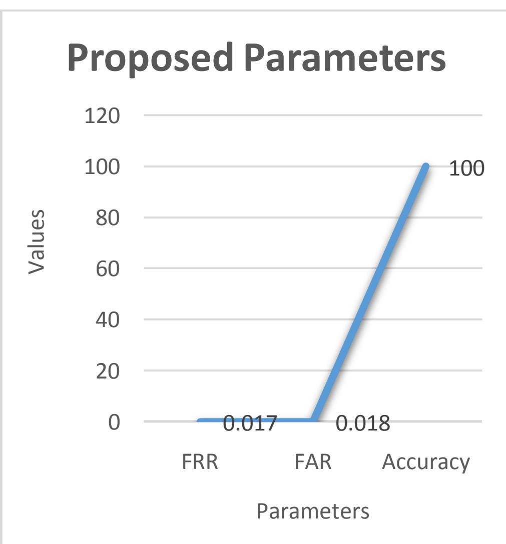


Figure: 8Proposed Parameters

Table 1 Proposed Parameter Values

Parameter	Values
FRR	.017
FAR	.018
Accuracy	100

6. CONCLUSION AND FUTURE WORK

In this paper the problem of music classification will be solved using combination of MFCC and NN.

Firstly feature extraction will be done using MFCC. Then optimization of feature set will be done using genetic algorithm. Then, at first, the neural network is trained based on the features of music files in the database. The image features considered here are average value, min value, sampling value, pitch value and max. Value. The training is carried out using NN algorithm. This trained when presented with then similar to feature extracted music files will be displayed or recognized from the database i.e. jazz, hip-hop, pop, classical and rock. The results show a considerable improvement in terms of FRR = .017, FAR = .018 and accuracy =100% of music classification rate.

This work can be extended by integrating with Fuzzy C-means clustering algorithm for better efficiency.

REFERENCES

[1] G. Agostini, M. Longari, and E. Poolastri. Musical instrument timbres classification with spectral features. *EURASIP Journal Appl. Signal Process*., vol. 1, **(2003)**, pp. 5–14.

[2] J. C. Brown, O. Houix, and S. McAdams. Feature dependence in the automatic identification of musical woodwind instruments. *Journal of Acoustic Soc. America*, vol. 109, issue 3, **(2001)**, pp. 1064–1072.

[3] S. Hao Chen, S. Huang Chen, and R. C. Guido. Music genre classification algorithm based on dynamic frame

analysis and support vector machine. In *IEEE Intl. Symposium on Multimedia*, **(2010)**, pp. 357–361.

[4] Y. M. G. Costa, L. S. Oliveira, A. L. Koreich, and F. Gouyon. Music genre recognition using spectrograms. In *Intl. Conf. on Systems, Signal and Image Processing*, **(2011)**, pp. 1–4.

[5] J. D. Deng, C. Simmermacher, and S. Cranefield. A study on feature analysis for musical instrument classification. *IEEE Trans. on System, Man and Cybernatics*– Part B, vol. 38, no. 2, **(2008)**, pp. 429–438.

[6] W. J. Dowling and D. L. Harwood. *Music Cognition*. Academic Press, Inc., **(1986)**.

[7] J. Eggink and G. J. Brown. Instrument recognition in accompanied sonatas and concertos. In *IEEE Conf. on Acoustics, Speech and Signal Processing*, **(2004)**, pp. 217–220.

[8] A. Eronen. Comparison of features for musical instrument recognition. In *IEEE Workshop Appl. Signal Process. Audio Acoust.*,**(2001)**, pp. 19–22.

[9] F. Fernandez, F. Chavez, R. Alcala, and F. Herrera. Musical genre classification by means of fuzzy rule-based systems: A preliminary approach. In *IEEE Congress on Evolutionery Computing*, **(2011)**, pp. 2571–2577.

[10] McKay, Cory. Automatic music classification with jMIR. *Dis*s. McGill University, **2010.**

[11] Ju, Han, Jian-Xin Xu, and Antonius MJ VanDongen. "Classification of musical styles using liquid state machines." Neural Networks (IJCNN), *The 2010 International Joint Conference on*. *IEEE*, **2010**.

[12] Nasridinov, Aziz, and Young-Ho Park. "A Study on Music Genre Recognition and Classification Techniques." *International Journal of Multimedia & Ubiquitous Engineering* 9.4 **(2014).**

[13] Shaveta Sharma, Parminder Singh" Extracting GFCC Features for Emotion Recognition from Audio Speech Signals*"* *International Journal of Advanced Research in Computer Science and Software Engineering*,Volume 5, Issue 1, January **2015.**