IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AANN based Music Genre Classification using PLP

[1]R. Thiruvengatanadhan, [2]P. Kathirvel, [3]R. Seenu

[1]Assistant Professor, [2]Assistant Professor/Programmer, [3]Assistant Professor/Programmer
[1]Department of Computer Science and Engineering
Annamlai University, Annamalainagar, Tamilnadu, India
Deputation (Government Polytechnic College, Arakandanallur, Villupuram)

[2]Department of Computer and Information Science
Annamlai University, Annamalainagar, Tamilnadu, India
Deputation (Government Polytechnic College, Arakandanallur, Villupuram)

[3]Department of Computer and Information Science
Annamlai University, Annamalainagar, Tamilnadu, India
Deputation (Arignar Anna Government Arts College, Villupuram)

Abstract: Automatic music genre classification is very useful in music indexing. In audio indexing, automatic music classification is more useful content based music retrieval and online audio distribution. Perceptual Linear Prediction (PLP) features are extracted from the music signal. This paper describes a technique that uses Autoassociative Neural Network (AANN) to classify music. The proposed feature extraction and classification models results in better accuracy in music genre classification.

Index Terms **- Music Signal, Feature Extraction, Perceptual Linear Prediction (PLP), Autoassociative Neural Network (AANN).**

I. Introduction

Speech, music, and any other sound source, as well as their combination, are all considered audio. File name, file format, sample rate, and other parameters make up audio. The need to naturally arrange, to which class a sound has a place, makes sound order and classification an arising and significant exploration region [1]. Machine learning has become very popular in recent years. Depending on the type of application and the data set available, certain types of machine learning techniques are more appropriate than others for different applications [2]. Genre classification is the way that can classify similar types of data into a single identity and give that identity as its name. The classification of such songs every day will become a tired activity, where the technology can be used to cure the music and make classification easier or more efficient by using its rhythms, beats and lyrical composition [3]. During the recent years, there have been many studies on automatic audio classification using several features and techniques. Due to improvements in internet services and network bandwidth there is also an increase in number of people involving with the audio libraries. Music has additionally been isolated into Genres and sub kinds on the premise on music as well as on the verses too [4]. This makes order more enthusiastically. To cause things more to confound the meaning of music type may have very much changed over the long run [5]. For instance, rock songs that were made fifty years ago are different from the rock songs we have today.

II. PERCEPTUAL LINEAR PREDICTION (PLP)

Hermansky developed a model known as PLP. It is based on the concept of psychophysics theory and discards unwanted information from the human pitch [6]. PLP is an approximation of three components of the perceptron, namely the crucial band resolution curves, the equal loudness curve, and the power law relation of intensity loudness. PLP coefficients are often used because they approximate well the high-energy regions of the speech spectrum while simultaneously smoothing out the fine harmonic structure, which is often characteristic of the individual but not of the underlying linguistic unit [7].

The process of PLP computation is shown in Figure 1. The audio signal is hamming windowed to reduce discontinuities. The Fast Fourier Transform (FFT) transforms the windowed speech segment into the frequency domain [8]. The auditory warped spectrum is convolved with the power spectrum of the simulated critical-band masking curve to simulate the critical-band integration of human hearing. The frequency bandwidth formed by the cochlea, which works as an auditory filter, is known as the critical band. Bark scale corresponds to 1 to 24 critical bands.

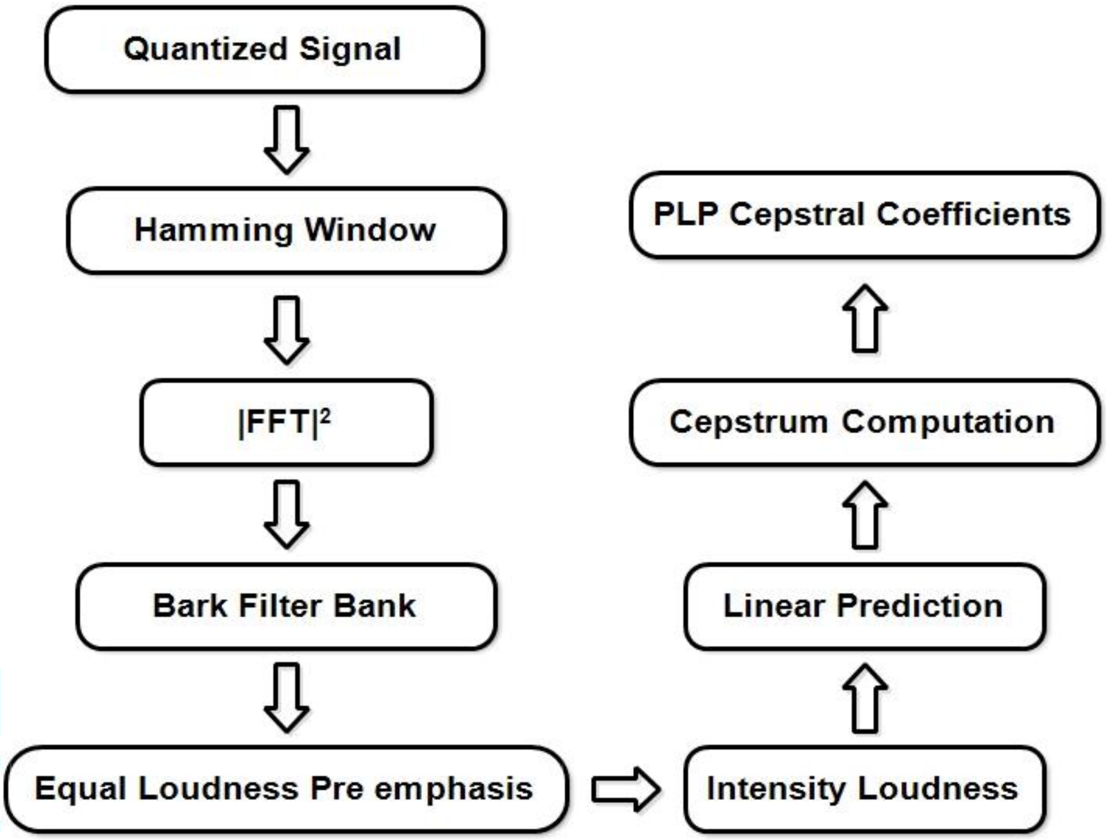

Fig.1 PLP Parameter Computations.

A weight function is added to the sampled values using an equal loudness curve to simulate the human hearing sensitivity at varying frequencies. The intensity loudness power law is an approximation of the power law of hearing, which relates sound intensity and perceived loudness of the sound [9]. Either the coefficients can be used as such for representing the signal or they can further be transformed to Cepstral coefficients. In this work, a 9th order LP analysis is used to approximate the spectral samples and hence obtained a 9-dimensional feature vector for a speech signal of frame size of 20 milliseconds is obtained.

III. Autoassociative Neural Network (AANN)

Autoassociative Neural Network (AANN) model comprises of five layer network which catches the dissemination of the component vector. The information layer in the organization has less number of units than the second and the fourth layers. The number of units in the first and fifth tiers is higher than in the third layer [10]. The quantity of preparing units in the subsequent layer can be either straight or non-direct. Yet, the handling units in the first and third layer are non-direct. Back proliferation calculation is utilized to prepare the organization [11]. The shape of the hyper surface is determined by projecting the cluster of feature vectors in the input space onto the lower dimensional space simultaneously, as the error between the actual and the desired output gets minimized. During testing the acoustic elements removed are given to the prepared model of AANN and the normal blunder is gotten. The structure of the AANN model used in our study is 9L 12N 3N 12N 9L for PLP, for capturing the distribution of the acoustic features.

IV. EXPERIMENTAL RESULTS

4.1 The database

The data is collected from Marsyas databases with distinct characteristics for genre classification. In our work dataset consists of 550 audio tracks each 30 seconds long. It contains 4 genres. The tracks are all 22050Hz Mono 16-bit audio files in .wav format.

Table 1 Classes and number of samples in the Marsyas database

Classes	Samples
Jazz	200
Pop	90
Metal	220
Disco	40

4.2 Acoustic feature extraction

In this work fixed length frames with duration of 20 ms and 50 percentages overlap (i.e., 10 ms) are used. The objective of overlapping neighbouring frames is to consider the harmonic information characteristic of audio content. An input wav file is given to the feature extraction techniques. PLP 9 dimensional feature values will be calculated for the given wav file. The above process is continued for 100 number of wav files.

4.3 Classification

The training process analyzes music training data to find an optimal way to classify music frames into their respective classes. The feature vectors are given as input and compared with the output to calculate the error. The performance of music classification is studied by varying the number of units in the compression layer as shown in Figure 2.

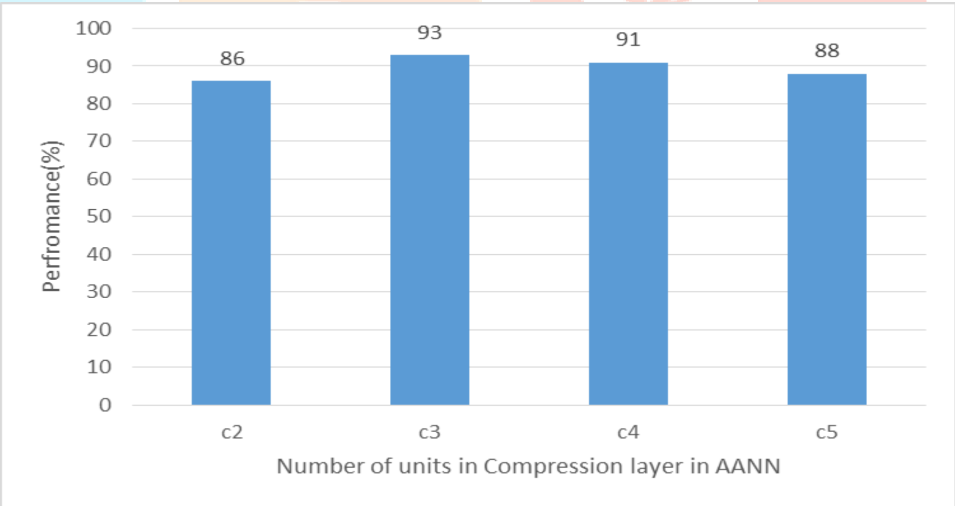

Fig. 2 Performance of Music Classification in Terms of Number of Units in the Compression Layer

The performance of speech recognition in terms of number of units in the expansion layer is shown in Figure 3. The network structures 9L 12N 3N 12N 9L gives a good performance and this structure is obtained after some trial and error.

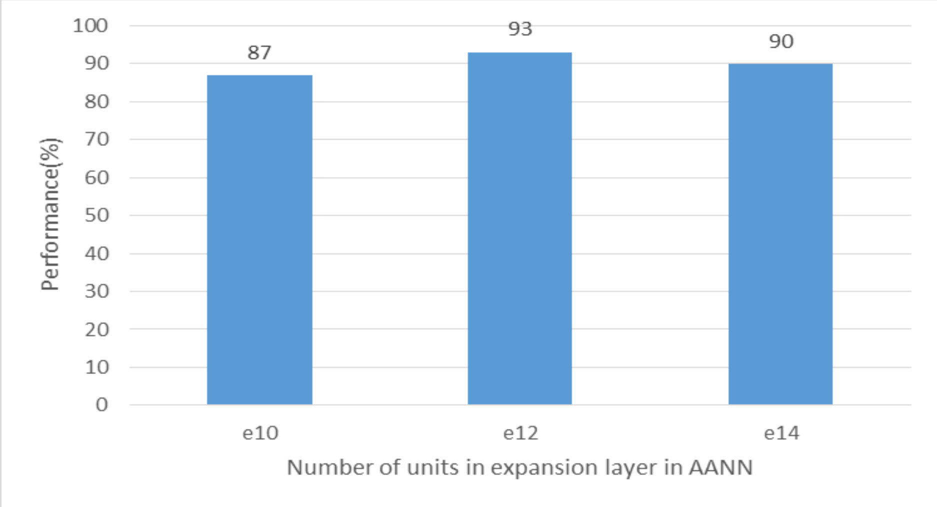

Fig. 3 Performance of Speech Recognition in Terms of Number of Units in the Expansion Layer.

IV. CONCLUSION

In this paper, we have proposed a programmed music sort arrangement framework utilizing AANN. PLP is calculated as features to characterize music content. AANN learning calculation has been utilized for the order of kind classes of music by gaining from preparing information. The certainty score is determined from the standardized squared blunder and the classification is concluded dependent on most noteworthy certainty score between Jazz and pop, pop and rock by learning from training data. It shows that the proposed method can achieve better classification accuracy than other approaches. Experimental results show that the proposed audio AANN method has good performance in musical genre classification scheme is very effective and the accuracy rate is 93%.

REFERENCES

[1] H Watanabe SM, Kikuchi H (2010) Interval calculation of em algorithm for gmm parameter estimation. Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium pp 2686–2689

[2] N. Pelchat and C. M. Gelowitz, "Neural Network Music Genre Classification," in Canadian Journal of Electrical and Computer Engineering, vol. 43, no. 3, pp. 170-173, Summer 2020, doi: 10.1109/CJECE.2020.2970144.

[3] A. Ghildiyal, K. Singh and S. Sharma, "Music Genre Classification using Machine Learning," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 1368-1372, doi: 10.1109/ICECA49313.2020.9297444.

[4] Vaishali Jabade, Vedang Deshpande and Aditya K Kumar. Music Generation and Song Popularity Prediction using Artificial Intelligence - An Overview. International Journal of Computer Applications, April (2019), 182(50):33-39.

[5] Dijk, L. Van. Radboud Universiteit Nijmegen Bachelorthesis Information Science Finding musical genresimilarity using machine learning techniques, (2014),1–25.

[6] Peter M. Grosche, Signal Processing Methods for Beat Tracking, Music Segmentation and Audio Retrieval, Thesis, Universit¨at des Saarlandes, 2012.

[7] Alif Noushad et al, International Journal of Computer Science and Mobile Applications, Vol.6 Issue. 2, February- 2018, pg. 131-138

[8] PetrMotlcek, Modeling of Spectra and Temporal Trajectories in Speech Processing, PhD thesis, Brno University of Technology, 2003.

[9] Poonam Sharma and Anjali Garg. Feature Extraction and Recognition of Hindi Spoken Words using Neural Networks. International Journal of Computer Applications 142(7):12-17, May 2016.

[10] D. Li, I. K. Sethi, N. Dimitrova, and T. Mc Gee, "Classification of General Audio Data for Content Based Retrieval," Pattern Recognition Letters, vol. 22, no. 1, pp. 533-544, 2001.

[11] N. Nitananda, M. Haseyama, and H. Kitajima, "Accurate Audio-Segment Classification using Feature Extraction Matrix," IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 261-264, 2005