International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 10 | Oct 2018 www.irjet.net p-ISSN: 2395-0072

Music Genre Classification using SVM

R. Thiruvengatanadhan[1]

[1]Assistant Professor/Lecturer (on Deputation), Department of Computer Science and Engineering, Annamalai University, Annamalainagar, Tamil Nadu, India

---

Abstract *:- Automatic music genre classification is very useful in music indexing. Tempogram is one of the feature extraction method uses in classification of musical genre that is based temporal structure of music signals. Searching and organizing are the main characteristics of the music genre classification system these days. This paper describes a new technique that uses support vector machines to classify songs. Support vector machines classify audio into their respective classes by learning from training data. The proposed feature extraction and classification models results in better accuracy in music genre classification.*

Key Words: Feature Extraction, Tempogram and Support vector machines (SVM).

1. INTRODUCTION

There are numerous studies that are investigated in the field of digital music and how it would be possible to enhance user's experience. A lot of untagged music files are being archived, while some contain assumed or false tags. However automatic genre classification is not an easy task considering music evolving within short periods. In addition to this, the advancement in digital signal processing and data mining techniques has led to intensive study on music signal analysis like , content-based music retrieval, music genre classification, duet analysis, Musical transcription , Musical Information retrieval and musical instrument detection and classification. Musical Instrument detection techniques have many potential applications such as detecting and analyzing solo passages, audio and video retrieval, music transcription, playlist generation, acoustic environment classification, video scene analysis and annotation etc [1]

Advanced music databases are continuously achieving reputation in relations to specialized archives and private sound collections. Due to improvements in internet services and network bandwidth there is also an increase in number of people involving with the audio libraries. But with large music database the warehouses require an exhausting and time consuming work, particularly when categorizing audio genre manually. Music has also been divided into Genres and sub genres not only on the basis on music but also on the lyrics as well [2]. This makes classification harder. To make things more complicate the definition of music genre may have very well changed over time [3]. For instance, rock songs that were made fifty years ago are different from the rock songs we have today.

2. ACOUSTIC FEATURES FOR AUDIO CLASSIFICATION

An important objective of extracting the features is to compress the speech signal to a vector that is representative of the meaningful information it is trying to characterize. In these works, music features namely Tempogram features are extracted.

2.1 Tempogram

An element which gives shape to the music in temporal dimension is the rhythm. Rhythmic feature arranges sounds and silences in time. A predominated pulse called beat which serves as basis for temporal structure of music is induced [242]. Tempogram captures the local tempo and beat characteristics of music signals. The Fourier tempograms are used in the research work.

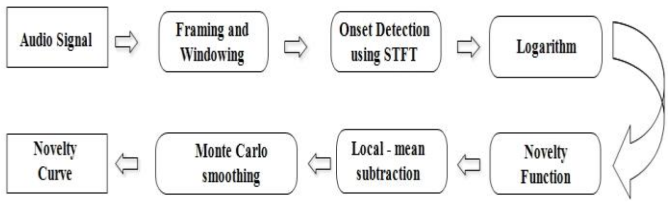

Fig -1: Novelty Curve Computations.

Human perceives rhythm as a regular pattern of pulses as a result of moments of musical stress. Abrupt changes in loudness, timbre and harmonic causes the occurrences of musical accents [4]. In instruments like piano, percussion instruments and guitar, occurs a sudden change in signal energy accompanied by very sharp attacks. A novelty curve is based on this observation and is computed for extracting meaningful information regarding note onset e.g. pieces of songs which are dominated by instruments [5]. In the pre-processing, stage short segmented frames have been extracted and windowed. The novelty curve computed, as described above, indicates peaks which represent note onset values [6]. A hamming window function is applied to avoid boundary problems as smoothing [7]. Novelty curve computation is shown in Fig. 1. The Fourier tempogram is calculated. Tempo related to musical context is a measure of beats per minute. Finally, the histogram is computed for each frame resulting in 12 dimensional feature vectors.

3. CLASSIFICATION MODEL

3.1 Support Vector Machine

A machine learning technique which is based on the principle of structure risk minimization is support vector machines. It has numerous applications in the area of pattern recognition [8]. SVM constructs linear model based upon support vectors in order to estimate decision function. If the training data are linearly separable, then SVM finds the optimal hyper plane that separates the data without error [9].

Fig. 2 shows an example of a non-linear mapping of SVM to construct an optimal hyper plane of separation. SVM maps the input patterns through a non-linear mapping into higher dimension feature space. For linearly separable data, a linear SVM is used to classify the data sets [10]. The patterns lying on the margins which are maximized are the support vectors.

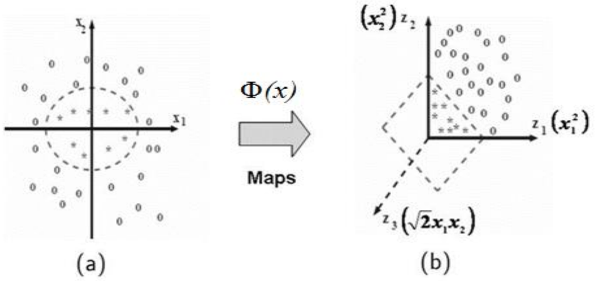

Fig -2: Example for SVM Kernel Function Φ(x) Maps 2-Dimensional Input Space to Higher 3-Dimensional Feature Space. (a) Nonlinear Problem. (b) Linear Problem.

The support vectors are the (transformed) training patterns and are equally close to hyperplane of separation. The support vectors are the training samples that define the optimal hyperplane and are the most difficult patterns to classify [11]. Informally speaking, they are the patterns most informative of the classification task. The kernel function generates the inner products to construct machines with different types of non-linear decision surfaces in the input space [12].

4. IMPLEMENTATION

4.1 Dataset Collection

The music data is collected from music channels using a TV tuner card. A total dataset of 100 different songs is recorded, which is sampled at 22 kHz and encoded by 16-bit. In order to make training results statistically significant, training data should be sufficient and cover various genres of music.

4.2 Feature Extraction

In this work fixed length frames with duration of 20 ms and 50 percentages overlap (i.e., 10 ms) are used. An input wav file is given to the feature extraction techniques. Tempogram 12 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

When the feature extraction process is done the music should be classified. We select 75 music samples as training data including 25 classic music, 25 pop music and 25 rock music. The rest 25 samples are used as a test set. For the SVM-1 which is used to classify music into pop and classic used for training. For the SVM-2 which is used to classify classic and rock are used for training. Table 1 shows Performance of music classification in different SVM kernel function.

Table -1: Performance of music genre classification in different SVM kernel function.

SVM Kernels	Performance
Polynomial	89%
Gaussian	95%
Sigmoidal	88%

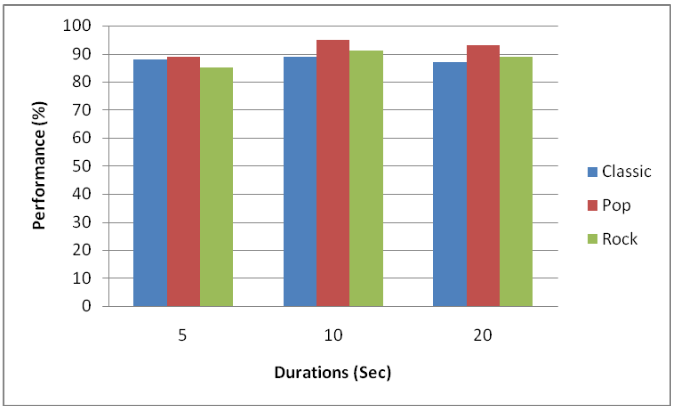

Chart -1: Performance of music classification for different duration of music clips

The performance of SVM for different duration as shown in Chart 1 shows that when the duration were increased from 10 to 20 there was no considerable increase in the performance.

5. CONCLUSION

In this paper, we have proposed an automatic music genre classification system using SVM. Tempogram is calculated as features to characterize audio content. SVM learning algorithm has been used for the classification of genre classes of music by learning from training data. Two nonlinear support vector machine classifiers are developed to obtain the optimal class boundaries between classic and pop, pop and rock by learning from training data. Experimental results show that the proposed audio support vector machine learning method has good performance in musical genre classification scheme is very effective and the accuracy rate is 95%.

REFERENCES

[1] C. Joder, S. Essid, G. Richard, and S. Member, "Temporal Integration for Audio Classification With Application to Musical Instrument Classification," IEEE Transactions on Speech and Audio Processing, vol. 17, no. 1, pp. 174–186, 2009.

[2] Serwach, M., & Stasiak, B., GA-based parameterization and feature selection for automatic music genre recognition. In Proceedings of 2016 17th International Conference Computational Problems of Electrical Engineering, CPEE 2016.

[3] Dijk, L. Van., Radboud Universiteit Nijmegen Bachelorthesis Information Science Finding musical genre similarity using machine learning techniques, 1–25, 2014.

[4] Mi Tian, Gy ¨ orgy Fazekas, Dawn A. A. Black, Mark Sandler, On the use of the tempogram to describe audio content and its Application to music structural segmentation, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015.

[5] Venkatesh Kulkarni, Towards Automatic Audio Segmentation of Indian Carnatic Music, Master Thesis, Friedrich Alexander University, 2014.

[6] Eronen, A. and Klapuri, A., "Music Tempo Estimation with k-NN regression," IEEE Transactions on Audio, Speech and Language Processing, vol. 18, no. 1, pp. 50-57, 2010.

[7] Gui, Wenming & Sun, Yao & Tao, Yuting & Li, Yanping & Meng, Lun & Zhang, Jinglan., A Novel Tempogram Generating Algorithm Based on Matching Pursuit. Applied Sciences, 2018.

[8] Chungsoo Lim Mokpo, Yeon-Woo Lee, and Joon-Hyuk Chang, "New Techniques for Improving the practicality of a SVM-Based Speech/Music Classifier," IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1657-1660, 2012.

[9] Hongchen Jiang, JunmeiBai, Shuwu Zhang, and Bo Xu, "SVM-Based Audio Scene Classification," IEEE International Conference Natural Language Processing and Knowledge Engineering, Wuhan, China, pp. 131-136, October 2005.

[10] Lim and Chang, "Enhancing Support Vector Machine-Based Speech/Music Classification using Conditional Maximum a Posteriori Criterion," Signal Processing, IET, vol. 6, no. 4, pp. 335-340, 2012.

[11] Md. Al Mehedi Hasan and Shamim Ahmad. predSucc-Site: Lysine Succinylation Sites Prediction in Proteins by using Support Vector Machine and Resolving Data Imbalance Issue. International Journal of Computer Applications 182(15):8-13, September 2018.

[12] Hend Ab. ELLaban, A A Ewees and Elsaeed E AbdElrazek. A Real-Time System for Facial Expression Recognition using Support Vector Machines and k-Nearest Neighbor Classifier. International Journal of Computer Applications 159(8):23-29, February 2017.