*Impact Factor (SJIF): 3.632*

**International Journal of Advance Research in Engineering, Science & Technology**

*e-ISSN: 2393-9877, p-ISSN: 2394-2444*

***(Special Issue for ITECE 2016)***

# Present Methodologies for Audio Base Music Classification

**Vinod R Patil[1], Manisha S Patil[2], Tushar H Jaware[3]**

[1]*Asso Prof, Dept of E&TC, RCPIT Shirpur*
[2] *Asst Prof, Dept of Computer Engg, SSVP's BSD COE Dhule*
[3]*Research Scholar, SSGMCE Shegaon*

*Abstract* ***—Large audio database available like storage device , CDs, internet, etc. music information retrieval an emerging research area .There are many classification class in music like genre classification, mood classification, artist recognition, instrument recognition, etc. In this paper we summarize the different types of feature set and techniques for feature extraction and music classification. The paper focuses on different types of classifier used for different classification tasks.***

***Keywords- feature selection, feature extraction, classifier, MIR, audio, genre.***

## I. INTRODUCTION

Music information retrieval (MIR) is an emerging research area in multimedia with the development of information and multimedia technologies, digital music data has become widely avail-able from various way, including radio broadcasting, digital storage such as compact discs (CDs), the Internet, etc. need to developing tools to effectively and efficiently retrieve and manage the music of interest to the end users

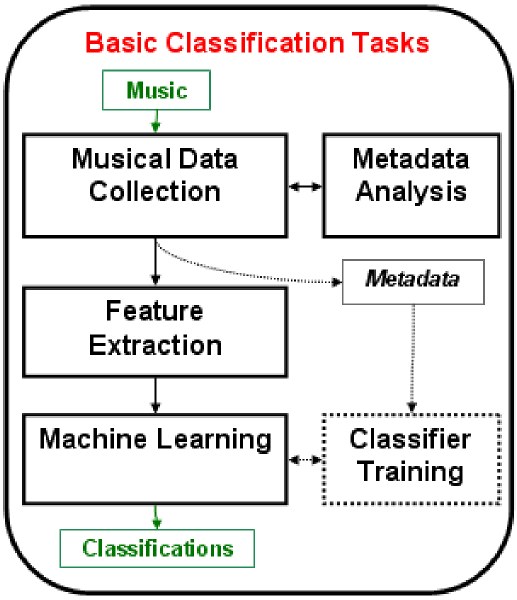


***Figure 1 Graphical Overview of Music Classification Tasks***

Music classification is a hot topic with many potential applications. It provides important functionalities for music retrieval. Many users may only be interested in specific types of music. Thus, a classification system would enable them to search for the music they are interested in. different types of music has different propriety. So manage them more effectively and efficiently

All Rights Reserved, @IJAREST-2016

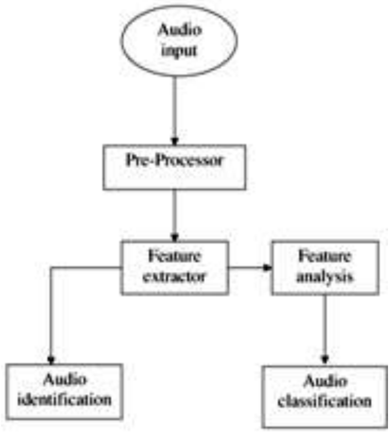


***Figure 2. Generalized music classification system***

Music classification has received much attention from MIR researchers in recent years. Different tasks are related to music classification are listed in the following.

- Genre Classification
- Mood Classification
- Artist Identification
- Instrument Recognition

In this paper, we provide an overview of features and techniques used for the music classification tasks. [26], [27] this article, review the state-of-the art in automatic genre classification of music collections. Different features and types of classifiers being developed and used.

## II. AUDIO FEATURES

Audio feature the key components of a music classification system. Different types of feature have categorized the into four subcategories, namely short-term features, long-term features, semantic features, and compositional features Scaringella [26]

| Class | Feature type | Used in |
|---|---|---|
| **Timber** | Zero crossing rate (ZCR) | [2], [3], [8], [10] |
| | Spectral Centroid (SC) | [2], [3] ,[8], [10],[25] |
| | Spectral Rolloff (SR) | [2], [3] ,[8], [10],[25] |
| | Spectral flux (SF)) | [2], [3] ,[25] |
| | Spectral bandwidth (SB) | [2] ,[8], [10], [25] |
| | Amplitude spectrum envelop (ASE) | [29],[14] |
| | Mel-frequency Cepstrum Coefficient (MFCC) | [2],[21],[8],[22] |
| | Fourier Cepstrum coefficient | [2] |
| **Temporal** | Amplitude modulation | [30],[31],[9],[14] |
| | Auto regressive modeling | [4],[13] |

***Table. 1Summary of Low-Level Feature Used In Music Classification***

All Rights Reserved, @IJAREST-2016

Audio features are divided into two levels, low-level and mid-level features, Low-level features can be further divided into two classes of timbre and temporal features. Timbre features capture the tonal quality of sound that is related to different instrumentation [1], low-level features again divided into two type of timbre and temporal features. Timbre features capture the tonal quality of sound that is related to different instrumentation, whereas temporal features capture the variation and evolution of timbre over time. Low-level features are obtained directly from various signal processing techniques like Fourier transform, spectral/cepstral analysis, autoregressive modeling, etc. mid-level features provide a closer relationship and include mainly three classes of features, namely rhythm, pitch, and harmony.

The large number of timbre features proposed for music classification such as spectral flux, spectral bandwidth, short time Fourier transform, one of the most powerful features such as MFCC[10]. Low level feature use to genre and mood classification [3], [32] .Low-level features has good performance for music classification, but mid-level features use for cover song detection [91], [93], mid-level features use for music analysis include Rhythm, Pitch, Harmony. Rhythm is the most widely used mid-level feature in audio-based music classification. Tzanetakis proposed the use of beat histogram (BH) [2] for genre classification. Rhythm feature use for mood classification [21], [22]. Pitch is one of the most useful features for music classification. Pitch feature can be used to determine fundamental frequency of sound. In Genre and mood classification pitch histogram use [1], [20].Rhythm features used in most of works for mood classification [2], [25], [27].

## III. MUSIC CLASSIFIER

There are different techniques for the music classification. The common classifiers are K-nearest neighbor (K-NN) [35], support vector machine (SVM) [36], and GMM classifier [28]. Different classifiers also used for different music classification, including logistic regression [10], [39], artificial neural networks (ANN) [4], [6], [25], [37], decision trees [5], [23] linear discriminate analysis (LDA) [2], [14], [42] nearest centroid (NC) [18], [20], and sparse representation-based classifier (SRC) [16]. For music classification K-NN [35] and SVM [36] are the most used, K-NN and SVM are related to single feature vector representations. Convolutional neural network (CNN) can directly handle feature. Feature learning is a process in which feature is automatically select and extract features for improving the classifications performance [4]. Feature selection/extraction can be done by supervised or unsupervised methods. The standard method for un-supervised feature extraction is principal component analysis (PCA). Large number features are available, and then combine them for music classification [10], [11], [22].

| feature | classifier | Accuracy % | Reference |
|---|---|---|---|
| **STFT+MFCC+Beat+Pitch** | KNN and GMM | **60** | G. Tzanetakis and P. Cook |
| **STFT+FET+MFCC+LPC** | Ada Boost.DT | **82** | J. Bergstra, N. Casagrande |
| **STFT+MFCC+Beat+Pitch** | SVM | **78** | T. Li and M. Ogihara |
| **CR+NTF** | SVM | **78** | I. Panagakis, E. Benetos |
| **MFSC+ASE+OSC** | NC | **90** | C.-H. Lin, J.-L. Shih, |
| **CR+NTF** | SRC | **92** | I. Panagakis, C. Kotropoulos |

***Table 2 Performance of Genre Classification Algorithms***

## IV. MUSIC CLASSIFICATION

Genre Classification: Genre classification is the wide area in MIR significant work by done by Tzane-takis and Cook [1], provide data set with 1000 songs in raw audio format evenly distributed in ten genres ISMIR 2004 also provide data set. Large work has been done with different approaches [2], [8], [1], and [12].

Mood classification: Mood classification is another important area. In this category song can be classify in to different emotional like happy, sad, and angry. if compare the feature and classification method between genre and mood classification lot of the features common with much more emphasis on low-level spectral features [10], [15], [18], rhythmic features have been used a lot in many mood classification systems [16], [17], [19].

Artist Identification: Another emerging area in music classification has artist identification it involves several subtasks—artist identification, singer recognition, and composer recognition. Since different artists, singers, and composers have their distinctive styles of performing, singing, and composition, it is possible to distinguish songs performed/sung/written by different artists/singers/composers based on the musical styles reflected in song audio, which is the purpose for general audio-based artist identification

All Rights Reserved, @IJAREST-2016

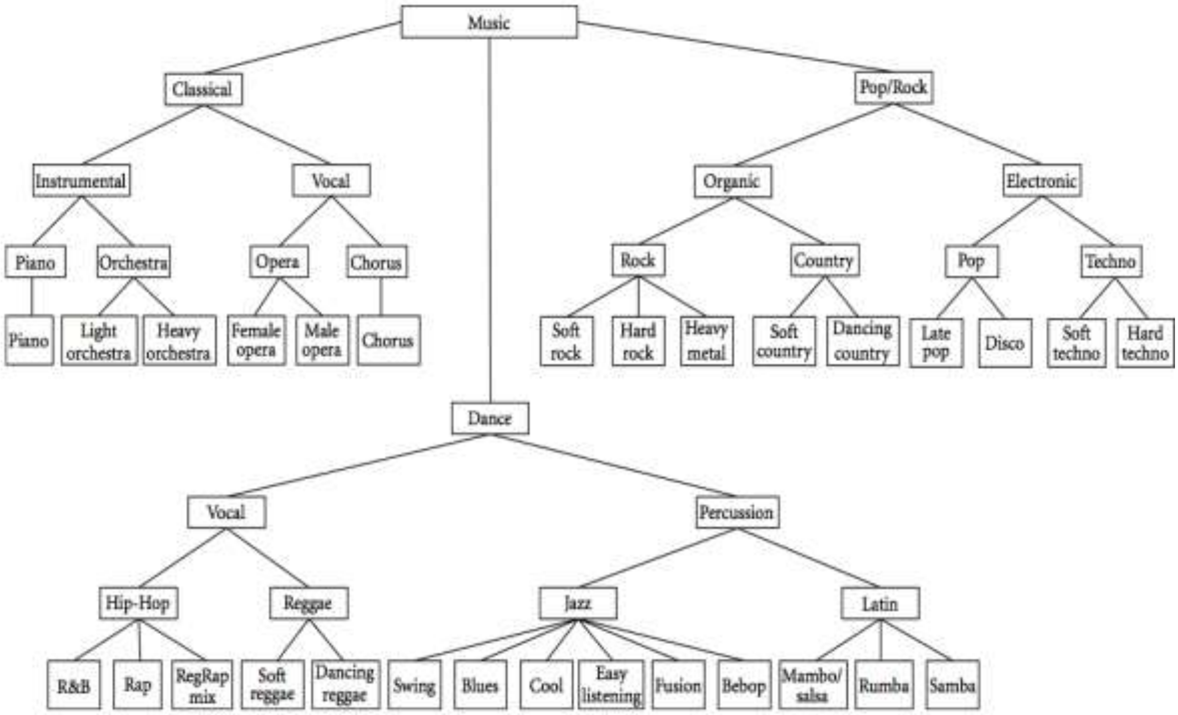


***Figure 3 Musical genre taxonomy***

Instrument Recognition: Instrument recognition is one another research area in music classification. The purpose is to identify the types of instruments playing at different intervals in the raw audio.

## V. CONCLUSION

In this paper review different types of feature and extraction technique and compared the results of their classification accuracy with different features as well as several classifiers. As can be observed, the music classification systems based on collaborating machine learning and evolutionary algorithms have better performance in comparison other methods.

## REFERENCES

[1] Zhouyu Fu, Guojun Lu, Kai Ming Ting, and Dengsheng Zhang “A Survey of Audio-Based MusicClassification and Annotation,”IEEETrans. Multimedia, vol. 13, no. 2, pp. 303–39, 2011

[2] G. Tzanetakis and P. Cook, “Musical genre classification of audio sig-nals,”IEEE Trans. Speech Audio Process., vol. 10, no. 5, pp. 293–302,2002.

[3] T. Li, M. Ogihara, and Q. Li, “A comparative study of content-based music genre classification,” inProc. SIGIR, 2003.

[4] T. Lidy and A. Rauber, “Evaluation of feature extractors and psycho-acoustic transformations for music genre classification,” inProc. Int. Conf. Music Information Retrieval, 2005.

[5] A. Meng and J. Shawe-Taylor, “An investigation of feature models for music genre classification using the support vector classifier,” inProc. Int. Conf. Music Information Retrieval, 2005.

[6] I. Mierswa and K. Morik, “Automatic feature extraction for classifying audio data,”Mach. Learn., vol. 58, pp. 127–149, 2005.

[7] N. Scaringella and G. Zoia, “On the modelling of time information for automatic genre recognition systems in audio signals,” inProc. Int. Conf. Music Information Retrieval, 2005.

[8] J. Bergstra, N. Casagrande, D. Erhan, D. Eck, and B. Kegl, “Aggregate features and ada boost for music classification,”Mach. Learn., vol. 65, no. 2–3, pp. 473–484, 2006.

[9] F. Morchen, A. Ultsch, M. Thies, and I. Lohken, “Modeling timbredistance with temporal statistics from polyphonic music,”IEEE Trans.Audio, Speech, Lang. Process., vol. 14, no. 1, pp. 81–90, 2006.

[10] T. Li and M. Ogihara, “Toward intelligent music information retrieval,”IEEE Trans. Multimedia, vol. 8, no. 3, pp. 564–574, 2006.

[11] J. Shen, J. Shepherd, and A. Ngu, “Towards effective content-based music retrieval with multiple acoustic feature combination,”IEEETrans. Multimedia, vol. 8, no. 6, pp. 1179–1189, 2006.

[12] T. Lidy, A. Rauber, A. Pertusa, and J. Inesta, “Improving genre clas-sification by combination of audio and symbolic descriptors using atranscription system,” inProc. Int. Conf. Music Information Retrieval,2007.

All Rights Reserved, @IJAREST-2016

[13] A. Meng, P. Ahrendt, and J. Larsen, “Temporal feature integration for music genre classification,” IEEE Trans. Audio, Speech, Lang.Process., vol. 15, no. 5, pp. 1654–1664, 2007.

[14] C.-H. Lin, J.-L. Shih, K.-M. Yu, and H.-S. Lin, “Automatic music genreclassification based on modulation spectral analysis of spectral and cep-stral features,”IEEE Trans. Multimedia, vol. 11, no. 4, pp. 670–682,2009.

[15] T. Li and M. Ogihara, “Detecting emotion in music,” inProc. Int. Conf.Music Information Retrieval, 2003.

[16] Y. Feng, Y. Zhuang, and Y. Pan, “Music retrieval by detecting mood via computational media aesthetics,” inProc. Int. Conf. Web Intelli-gence, 2003.

[17] D. Yang and W. Lee, “Disambiguating music emotion using software agents,” inProc. Int. Conf. Music Information Retrieval, 2003.

[18] L. Lu, D. Liu, and H.-J. Zhang, “Automatic mood detection and tracking of music audio signals,”IEEE Trans. Speech Audio Process., vol. 14, no. 1, pp. 5–18, 2006.

[19] Y.-H. Yang, Y.-C. Lin, Y.-F. Su, and H. H. Chen, “A regression ap-proach to music emotion recognition,”IEEE Trans. Audio, Speech,Lang. Process., vol. 16, no. 2, pp. 448–457, 2008.

[20] L. Mion and G. D. Poli, “Score-independent audio features for descrip-tion of music expression,”IEEE Trans. Audio, Speech, Lang. Process.,vol. 16, no. 2, pp. 458–466, 2008.

[21] M. Mandel and D. Ellis, “Song-level features and SVMs for music clas-sification,” inProc. Int. Conf. Music Information Retrieval, 2005.

[22] J. Shen, J. Shepherd, B. Cui, and K.-L. Tan, “A novel framework for efficient automated singer identification in large music databases,”ACMTrans. Inf. Syst., vol. 27, no. 3, pp. 1–31, 2009.

[23] S. Essid, G. Richard, and B. David, “Instrument recognition in poly-phonic music based on automatic taxonomies,”IEEE Trans. Audio, Speech, Lang. Process., vol. 14, no. 1, pp. 68–80, 2006.

[24] D. Little and B. Pardo, “Learning musical instruments from mixtures of audio with weak labels,” inProc. Int. Conf. Music Information Re-trieval, 2008.

[25] P. Hamel, S. Wood, and D. Eck, “Automatic identification of instrument classes in polyphonic and poly-instrument audio,” in Proc. Int. Conf. Music Information Retrieval, 2009.

[26] N. Scaringella, G. Zoia, and D. Mlynek, “Automatic genre classifica-tion of music content—A survey,”IEEE Signal Process. Mag., vol. 23,no. 2, pp. 133–141, 2006.

[27] M. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, and M.Slaney, “Content-based music information retrieval: Current direc-tions and future challenges,”Proc. IEEE, vol. 96, no. 4, pp. 668–696,2008.

[28] R. O. Duda and P. E. Hart, Pattern Classification, 2nd ed. New York:Wiley, 2000.

[29] H. G. Kim, N. Moreau, and T. Sikora, “Audio classification based on MPEG-7 spectral basis representation,”IEEE Trans. Circuits Syst.Video Technol., vol. 14, no. 5, pp. 716–725, 2004.

[30] E. Pampalk, A. Rauber, and D. Merkl, “Content-based organization and visualization of music archives,” inProc. ACM Multimedia, 2002.

[31] E. Pampalk, A. Flexer, and G. Widmer, “Improvements of audio-based music similarity and genre classification,” inProc. Int. Conf. Music Information Retrieval, 2005.

[32] B.-Y. Chua, “Automatic extraction of perceptual features and catego-rization of music emotional expression from polyphonic music audio signals,” Ph.D. dissertation, Gippsland School of Info. Tech., Monash Univ., Churchill, Australia, 2007.

[33] E. Gomez and P. Herrera, “The song remains the same: Identifyingversions of the same piece using tonal descriptors,” inProc. Int. Conf.Music Information Retrieval, 2006.

[34] J. Serra, E. Gomez, P. Herrera, and X. Serra, “Chroma binary similarity and local alignment applied to cover song identification,”IEEE Trans.Audio, Speech, Lang. Process., vol. 16, no. 6, pp. 1138–1151, 2008.

[35] T. Cover and P. Hart, “Nearest neighbor pattern classification,”IEEE Trans. Inf. Theory, vol. 13, no. 1, pp. 21–27, 1967.

[36] B. E. Boser, I. Guyon, and V. Vapnik, “A training algorithm for op-timal margin classifiers,” inProc. ACM Conf. Computational Learning Theory, 1992, pp. 144–152.

[37] A. Berenzweig, B. Logan, D. Ellis, and B. Whitman, “A large-scale evaluation of acoustic and subjective music similarity measures,” inProc. Int. Conf. Music Information Retrieval, 2003.

[38] Z. Fu, G. Lu, K. Ting, and D. Zhang, “On feature combination for music classification,” in Proc. Int. Workshop Statistical Pattern Recognition, 2010.

All Rights Reserved, @IJAREST-2016