Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 2 (2023)

Indian Film Songs Genre Wise Classification Based on Spectral Features

Miss. Darshita S. Pathak[1], Dr Tejas P. Patalia[2], Mr. Chitrang K. Vyas[3]

[1,]Lecturer, A.V. Parekh Technical Institute, Research Scholar, Gujarat Technological University, Ahmedabad Gujarat India,

[2]HOD Professor, V.V.P. Engineering College. PhD Supervisor, Gujarat Technological University, Ahmedabad Gujarat India,

[3]Lecturer, A.V. Parekh Technical Institute, Gujarat India,

***Abstract*:** The classification of musical genres has received a lot of attention in recent years due to the several uses it can serve in music information retrieval. Additionally, since genres frequently co-occur in music tracks in Indian Film Songs, it is desirable to record and model the genre correlations in order to enhance the effectiveness of multi-label music genre classification. Present a novel multi-modal strategy that uses a music spectral feature-based classification method to address these problems. In addition, a genre correlations extraction module is offered to identify and model probable genre correlations based on the characteristics of the proposed approach.

Index Terms: *Indian Music, Genre classification, spectral features, Indian filmy songs*

1. Introduction

Songs are an integral part of movies. Songs represent, enhance, and complement an emotion in the movie. In Indian Film industry which is widely accepted as word Bollywood exemplify movie songs. Indian Film Industry music one of such huge part of human lives in India and with rapid growth of Industrial collaboration over world .The versatility of musicality in the Hindi music industry. From dance music to orchestrated musicals, the industry serves you the multiple genre music at its epitome. Lyricists, producers, singers and artists are involved with this entertainment industry. Whether person is on car, auto-rickshaw or in radio, with wedding, sangeet ceremony or in funeral ceremony songs plays vital role as play by different genres. Classification of songs into different parts based on music spectral features involve music information retrieval algorithm for automatic categorization of songs.

2. Related Work

The interdisciplinary discipline of extracting information from music is known as music information retrieval (MIR). MIR is a young, expanding field of study with numerous practical applications. MIR practitioners may have training in signal processing, psychology, academic music research, musicology, machine learning, or a mix of these fields. [1]. Music signal feature extraction includes sound quality feature, melody feature and rhythm feature. “Genre” "kind" or "sort" is the term for any category of literature or other forms of art or entertainment.

[1]The singer's voice is incorporated with a variety of musical instruments in the music genre. Songs from the Jazz, Pop, Tarana, Sufi, Club, Hip-Hop, Disco, Filmy Sugam, Vocal Soul, and Fusion genres are frequently heard in Hindi film industries. The musical styles, the cultural setting, the themes' content, and their overall spirit can all be used to identify a music genre or subgenre. Although a single geographical category will frequently comprise a wide range of subgenres, geographical origin is occasionally used to designate a musical genre. According to Timothy Laurie, "genre has graduated from being a subset of popular music studies to being an almost ubiquitous framework for constituting and evaluating musical research objects" since the early 1980s. [2]
In Bollywood music Described genres can be the perfect example to justify the statement of accepting the western music culture, Names of given genres are classified in relation to Indian music by using abbreviation like Indo/Indie (India/Indian), many people also prefer to use the term ‘fusion’ music as well.

Music Genres

Filmy

Filmy ("of films") soundtracks are pieces of music composed and performed specifically for Indian cinema and the country's major motion picture industry. The majority of composers for motion pictures are music directors; playback singers provide the songs, and the genre accounts for 72% of India's music sales market. [7]

Romantic

Romantic composers sought to create music that was individualistic, emotional, dramatic and often programmatic; reflecting broader trends within the movements of Romantic literature, poetry, art, and philosophy. Romantic music was often ostensibly inspired by (or else sought to evoke) non-musical stimuli, such as nature, literature, poetry, super-natural elements or the fine arts. [8]

Sad

Blues is a kind of music that is renowned for being sorrowful. The blues song writers' difficult life are frequently depicted in the lyrics. [9]

Dance

The type of music you typically hear at bars and clubs is known as dance music. It is excellent for dancing to because it has a strong beat and is repetitive frequently. Dance music comes in a wide variety of styles, just like rock. They all will get your feet moving, regardless of how quickly or slowly they move. [9]

Remix

In modern dance music, remixes have become the norm, enabling a single song to be enjoyed in a variety of musical contexts or dancing environments. Such remixes frequently include new vocalists or musicians to the original mix as "featured" artists. Rap and hip hop music frequently employ the remix.

- Happy

A mid-tempo soul and neo soul tune, Happy. Songs that make people happy.

- Philosophical

Songs with philosophical themes include, but are not limited to, stories about the function and role of society, the meaning of life, ethics, morals, and the place of art in human lives, the value of experience, and the growth of knowledge.

- Rock

In the rock genre, traditional Indian instruments like the table, sitar, and dole are blended with electric guitar.

- Pop

Indian pop music that is danceable is referred to be "Indi pop," which is most often associated with Bollywood films. It was referred to as "sugar-coated dance melodies" and had Hindi lyrics and large orchestral film music. Indians appear to be more devoted to songs than performers.

- Party

Party is block-rocking, bass-driven hip-hop that only thinks about maintaining the groove. All of the lyrics are unimportant, lacking the political undertones of hardcore rap and only a small portion of the ingenuity of old school rap. Instead, the emphasis is on the song as a whole, with the bass and percussion taking center stage.

Spectral Audio Features

In the past three decades, as machine learning and deep learning have advanced, many experiments have been carried out to determine the most efficient methods for extracting numerical features from audio files as well as the top-performing classification algorithms for audio classification. Information retrieval systems for music is a suitable term for this specialized topic. Mel Frequency Cepstral Coefficients (MFCC) are one of the content-based acoustic features extraction methods from audio that Tao Li et al. identified in a comparative study [15], and the researchers also suggested a novel technique called Daubechies Wavelet Coefficient Histograms (DWCH). The requirement for a novel method proposal is established by the authors' observation that conventional feature extraction techniques largely capture partial information of musical signals. The researchers compared other[15], gave a thorough analysis of the most current audio approaches (as of 2016), classifying them first into the Physical and Perceptual categories, then into seven different types based on domain. Mel Frequency Cepstral Coefficients (MFCC) were classified as Wavelet Based Perceptual Features in the study's analysis of audio feature extraction methods. The methods for extracting musical features have developed further. On a set of unrelated music regression and classification tasks, K Choi conducted a comparison analysis between the traditional MFCC feature extraction technique and a pre-trained convnet feature. He discovered that the convnet feature outperformed the baseline MFCC feature in all the tasks taken into consideration [16], providing a foundation and justification for the success of the application of the transfer learning concept on music classification tasks.

A) MFCC

The MFCC is a representation defined as the real cepstrum of a windowed short-time signal derived from the fast Fourier transform of the speech signal. The use of a nonlinear frequency scale, which simulates the behavior of the auditory system, distinguishes this simulation from the actual cepstrum. The MFCC might look like this: [6]

MFCCs are commonly used as features in speech recognition systems, such as the systems which can automatically recognize numbers spoken into a telephone. The usage of MFCCs in applications for music information retrieval, such as genre categorization, audio similarity metrics, etc., is also on the rise.

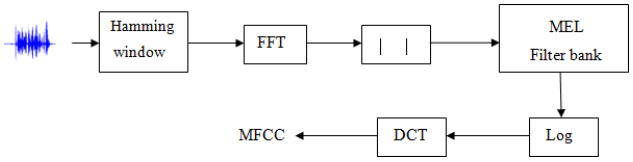

Figure 1. MFCC Calculation [6]

B) ZCR

An audio frame's Zero-Crossing Rate (ZCR) measures how frequently the signal's sign shifts across the frame. In other words, it is determined by dividing the number of times the signal's value shifts from positive to negative to vice versa by the frames duration. The ZCR is defined according to the following equation:

$$z_i = \sum_{m=1}^{n-1} sign[x_i(m-1)x_i(m)]$$

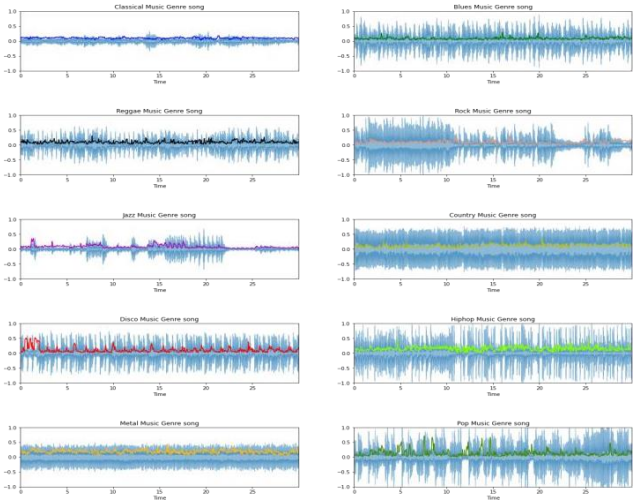

Figure 2. Sample ZCR values of different Genres

C) Centroid

A measurement used in digital signal processing to describe a spectrum is called the spectral centroid. It reveals the location of the spectral center of mass. It is strongly related to the perception of brightness in a sound perceptually. [10] It is often referred to as the spectral mass center. [11]It is determined using a Fourier transform to determine the weighted mean of the frequencies present in the signal, with their magnitudes serving as the weights. [12]

$$Centroid = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)}$$

Where x(n) represents the weighted frequency value, or magnitude, of bin number n, and f(n) represents the center frequency of that bin.

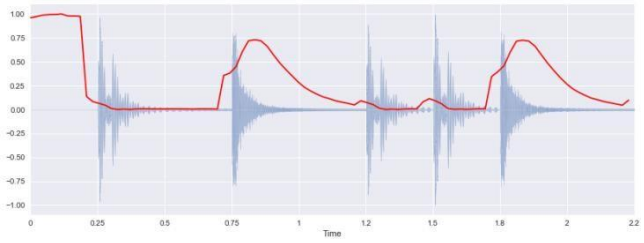

Figure 3. Centroid Value Signal Map

D) Roll Off

Specifically in electrical network analysis, and specifically in conjunction with filter circuits in the transition between a passband and a stopband, roll-off is the steepness of a transfer function with frequency. Although it is most frequently used to describe the network's insertion loss, it can theoretically be used to describe any significant function of frequency and any technology, not just electronics. Typically, roll-off is calculated as a function of logarithmic frequency.

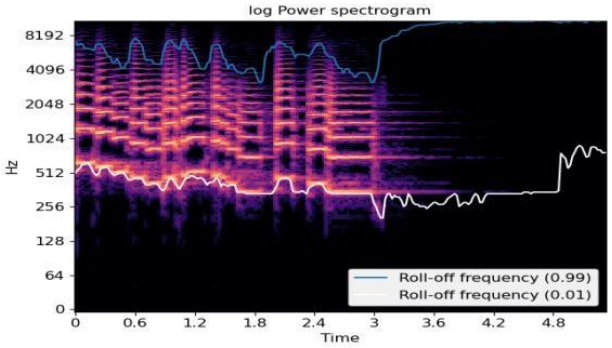

Figure 4. Log Power Spectrogram

Classification Algorithms

Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem. Maximum-likelihood training can be done by evaluating a closed-form expression. [17]

SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. [18]

K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm. [19]

Dataset

Dataset is prepared from hindigeetmala.com website retrieval using proposed algorithm shows results statistics of data as mentioned in table, dataset is prepared by generating retrieval approach by tagging mechanism of song title searching technique[13]. Computation time of retrieval of 100 songs contains almost Time : 4 minutes 17 sec and of 5076 Songs it takes Time Complexity measure as 2 hours 17 minutes (based on Internet availability and

speed) and Size of dataset is : 73,93,647 bytes. (Sample Size 7.4 MB on disk).

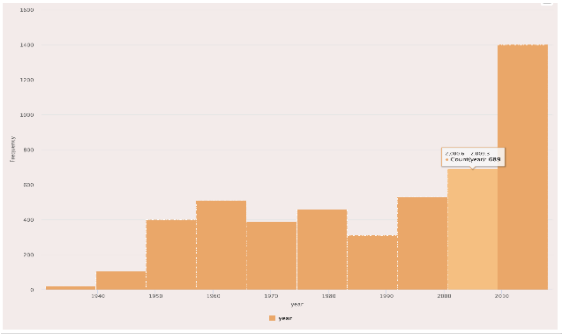

Chart 1. Year wise collected Song Statistics

No. of Songs	5076
No. of Lyricist	715
No. of Movies	2728
No. of Artist	1987
No. of Genres	57
Year Range	1940 - 2018

Table 1.Dataset Overall Statistics

Sanchay dataset is useful as training model for those who want to classify songs based on genre as well as Lyricist wise as it covers almost all years of Bollywood Hindi songs. Total of 57 genre are enlisted in retrieval process but in India 125 types of genres are available some of top genre are enlisted as below. Result shows Bollywood filmy songs contain most of Romantic and Dance number songs, while others are proceed as Sad, Philosophical, Dard bhare, Masti bhare etc.

Proposed Architecture

The technique presented over here is using kNN training algorithm model and testing as well. Here in WAV format audio files are given which are then pass through a set of steps on which various operation such as Labeling, Feature extraction then preprocessing training and Multi model training is implemented through which final prediction of genre of song is carried out as label genre of song.

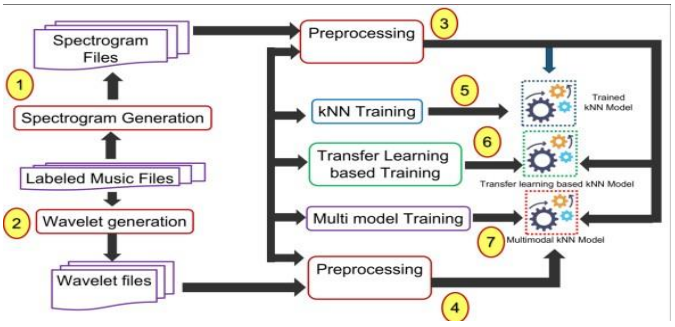

Figure 5.Proposed Architecture Training Testing Model

Step 1: The songs of 10 genres Classical, Tareef, Philosophical and Romantic etc. are considered.

Step 2: The database collected and sampled at a rate of 44100 Hz and 16 bit precision.

Step 3: 926 songs consisting of approx. 90 songs of each genre are considered in this research work. (10 Genres)

The music clips considered are downloaded from freely available online songs.

The training dataset consists of 917 songs and test data contains another 900 songs.

Spectral feature is used to differentiate the defined classes.

Training Model Structure

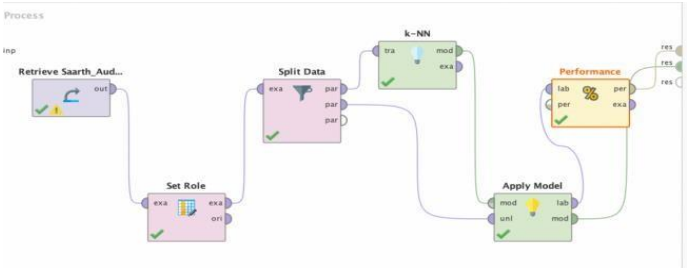

Figure 6. Training Process Model

Algorithm	Bangla Music	Hindi Music
SVM	0.7583	0.3283
NB	0.667	0.548
k-NN	0.6417	0.7163

Table 2: Accuracy of Algorithm Implementation on Bangla Dataset and Hindi Dataset

Accuracy Parameters

Many different measures for evaluating the performance of information retrieval systems have been proposed. The measures require a collection of documents and a query. All common measures described here assume a ground truth notion of relevancy: every document is known to be either relevant or non-relevant to a particular query. In practice queries may be ill-posed and there may be different shades of relevancy. Precision is the probability that a (randomly selected) retrieved document is relevant. Recall is the probability that a (randomly selected) relevant

document is retrieved in a search. F-measure is measure that combines precision and recall is the harmonic mean of precision and recall.

F= 2* ((precision* recall)/ (precision+ recall))

3. Results & Analysis

By Training the proposed model and calculation applied on 926 songs of dataset, the ranges of top 10 genres are shown in table. As perspective of instrumental structure of various genres table range is also shows that kind of tarana and Sufi which is having kind of same texture composition are in nearer range of each other thus they are as general resulted into Filmy category.

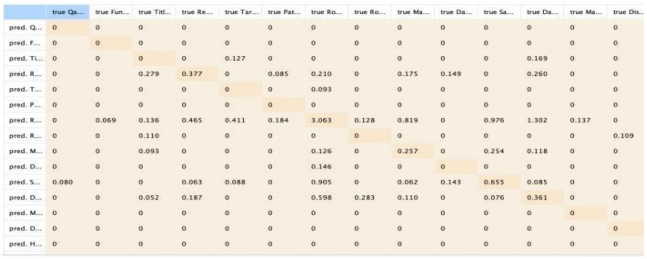

Figure 7. Confusion Matrix of Predicted Genres

	Accuracy(MFCC)	Accuracy(Roll Off)	Accuracy(Centroid)	Accuracy(ZCR)	Accuracy
Filmy	0.751	0.763	0.81	0.815	0.785
Romantic Songs	0.884	0.82	0.797	0.776	0.819
Sad Songs	0.717	0.717	0.774	0.811	0.755
Dance Songs	0.595	0.595	0.568	0.405	0.541
Remix Songs	0.63	0.63	0.778	0.704	0.686
Happy Songs	0.545	0.545	0.182	0.182	0.364
Philosophical Songs	0.182	0.182	0.273	0.364	0.25
Rock Songs	0.75	0.75	0.25	0.625	0.594
Pop Songs	0.5	0.5	0.333	0.333	0.417
Party Songs	0	0	0	0.25	0.063

Table 3. Accuracy of Genres

4. Conclusion

The proposed system drawn for this paper will be useful to fulfill its objective to make searching of song by its genre efficiently and also to classify track of song. The classification of music genres in this work is based on four spectral mode and genre correlations, and the results are shown by technique leveraging. This indicates that it is important to classify audio tracks based on genre in order to distinguish human emotion. On the Sanchay dataset,

the k-NN algorithm significantly outperforms the results of other algorithms and produces state-of-the-art results.

5. References

[1] Samson, Jim. "Genre". In Grove Music Online. Oxford Music Online. Accessed March 4, 2012

[2] Laurie, Timothy (2014). "Music Genre as Method". Cultural Studies Review. 20 (2). doi:10.5130/csr.v20i2.4149.

[3] "ID3 Tags " "d3v2.3.0 - ID3.org". id3.org.

[4] http://www.mpx.cz/mp3manager/tags.htm Archived 2013-10-07 at archive. Today - 404

[5] Mohan Rao B C*, Vinayak Arkachaari*, Harsha M N*, Sushmitha M N*, Gayathri Ramesh K K*, Ullas M S* ,Pathi Mohan Rao*, Sudha G*, Narayana Darapaneni#Dept of Computer Science & Engineering, PES University, Bengaluru, India. # Northwestern University SPS and Great Learning,Bi-Sampling Approach to Classify Music Mood leveraging RagaRasa Association in Indian Classical Music,IEEE Xplorer,2021

[6] Rea, C. & Macdonald, P. & Carnes, G.. (2010). Listening to classical,pop, and metal music: An investigation of mood. EMPORIA STATE RESEARCH STUDIES. 46. 1-3.

[7] Pinglay, Prachi (10 December 2009). "Plans to start India music awards". BBC News. Retrieved 2 May 2010.

[8] Truscott, Harold (1961). "Form in Romantic Music". Studies in Romanticism. 1 (1): 29–39. doi:10.2307/25599538. JSTOR 25599538.

[9] Description about "Music Genre s"https://englishlive.ef.com/blog/english-in-the-real-world/music-genres-english-favourite/

[10] Grey, John M.; Gordon, John W. (1978). "Perceptual effects of spectral modifications on musical timbres". The Journal of the Acoustical Society of America. Acoustical Society of America (ASA). 63 (5): 1493–1500. doi:10.1121/1.381843. ISSN 0001-4966.

[11] Pulavarti, Surya V. S. R. K.; Maguire, Jack B.; Yuen, Shirley; Harrison, Joseph S.; Griffin, Jermel; Premkumar, Lakshmanane; Esposito, Edward A.; Makhatadze, George I.; Garcia, Angel E.; Weiss, Thomas M.; Snell, Edward H. (2022-02-17). "From Protein Design to the Energy Landscape of a Cold Unfolding Protein". The Journal of Physical Chemistry B. 126 (6): 1212–1231. doi:10.1021/acs.jpcb.1c10750. ISSN 1520-6106. PMC 9281400. PMID 35128921.

[12] A Large Set of Audio Features for Sound Description - technical report published by IRCAM in 2003. Section 6.1.1 describes the spectral centroid.

[13] Folksonomy-based information retrieval by generating tag cloud for electronic resources management industries and suggestive mechanism for tagging using data mining techniques SD

[14] Pandya, PV Virparia - Web Usage Mining Techniques and Applications …, 2017

[15] Muhammad Salman Khan, Miao Yu , Pengming Feng , Liang Wang , Jonathon Chambers , "An unsupervised acoustic fall detection system using source separation for sound interference suppression", Signal Processing

[16] Volume 110, May 2015, Pages 199-210. ScienceDirect

[17] Sayf A. Majeed -Hadba University College , HAFIZAH HUSAIN , Salina Abdul Samad Universiti Kebangsaan Malaysia , Tarik Idbeaa -Jabal Al-Gharbi University ,Mel frequency cepstral coefficients (Mfcc)

feature extraction enhancement in the application of speech recognition: A comparison study. September 2015Journal of Theoretical and Applied Information Technology 79(1):38-56

[18] K. J. D'souza and Z. Ansari, "Big Data Science in Building Medical Data Classifier Using Naïve Bayes Model," 2018 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), Bangalore, India, 2018, pp. 76-80, doi: 10.1109/CCEM.2018.00020.

[19] Deldadeh Barani, Bahman & Daliri, Mohammad. (2014). Support Vector Machine (SVM) Classifier (پارسی Persian language). 10.13140/2.1.3276.0649.

[20] Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN COMPUT. SCI. 2, 160 (2021). https://doi.org/10.1007/s42979-021-00592-x