MUSIC EMOTION RECOGNITION FOR INDIAN FILM MUSIC

Makarand Velankar
Faculty CCOEW, PhD scholar PICT, SPPU Pune
makarand.velankar@cumminscollege.in

Dr. Parag Kulkarni
Chief Scientist, Kvinna Limited Pune
paragindia@gmail.com

EXTENDED ABSTRACT

Automatic music emotion recognition (MER) is a challenging problem considering different music dimensions. These dimensions include multimodal, multifaceted aspects of music on one side and the listener's perception on another side. Ground truth used generally is the human perception which has the influence of different factors such as cultural background, likes/dislikes, current emotional state, musical background, etc. Music emotional modeling is mainly influenced by the music genre for different emotion classes. Russel proposed a 2 dimensional model of Valence and arousal axes classifying emotions into 4 quadrants or cluster of emotions. As per ancient Indian Literature by Bharat Muni [1], the Natyasastra proposes navras or 9 states of mind as Shringar (romantic), hasya (happy), roudra (anger), karunya (sad), bibhastya (disgust), bhayanak (horrible), veer (bravery), adbhut(wonder), and shant (peace). These expressions are used widely in different classical dance performances in India. Anger, disgust, wonder and bravery are the less experienced emotions in case of Indian music. Bhakti (Devotional) is a different emotional class widely used in Indian music. Experimentation was proposed for automatic emotion classification for Indian music, which will be useful for automatic playlist generation or music search for different moods using content-based retrieval. The system used for the experiments is as shown in Figure 1

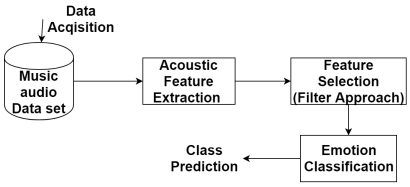

Figure 1. Music emotion recognition system

Data acquisition for preparing data set for different emotional classes was done from the study of popular music websites and streaming services in India. This study revealed that widely used emotional classes by Indian music lovers for play list or search purpose are happy, exiting or danceable, sad, romantic and devotional. Considering this, a data set of 500 random popular Indian songs in Hindi film music with 100 songs from each class was created for the experimentation for music emotion recognition using machine learning. Musical features with timbre, intensity, rhythm, spectral and melodic feature categories were extracted. Different feature selection approaches were used for identifying prominent discriminating features for MER. 4 different feature sets (FS1 to FS4) were selected with 72, 27, 17 and 28 features respectively. Different machine learning (ML) classifiers were used during the experiments as Naive Bayes (NB), Multilayer Perceptron (MLP), Support Vector Machine (SVM), J-48 Decision Tree (J48) and Random Forest (RF). Different machine learning classifiers provided the prediction accuracy in the range of 48.2 to 70 % for different feature sets as shown in Table 1. It was observed that accuracy depends on the classifier and the feature set both. Performance of MLP classifier looks promising compared to other classifiers.

ML Classifier	Feature Set (number of features) Accuracy in %			
	FS1 (72)	**FS2 (27)**	**FS3 (17)**	**FS4 (28)**
NB	54.4	51.2	48.4	53.8
MLP	60.2	**70**	64	63.8
SVM	64	62.6	59.2	62
J-48	53.2	53.6	49.2	54.8
RF	61.2	60.8	57.6	60.6

Table 1. Accuracy of ML classifiers for different feature sets

MLP classifier provided maximum accuracy of 70% for the feature set 2 (FS2), which was set of 27 distinct features. It indicates the use of distinct discriminating features provides better results. Confusion matrix with the predicted and actual class for this combination is as shown in Table 2. It can be observed from the confusion matrix that accuracy of the prediction for the romantic class is very low with confusion in prediction as sad and devotional. Exciting class emotions are well captured with 90 % accuracy.

	Romantic	**Exciting**	**Happy**	**Sad**	**Devotional**
Romantic	**43**	4	6	30	17
Exciting	0	**90**	8	2	0
Happy	3	10	**82**	4	1
Sad	20	4	2	**67**	7
Devotional	19	1	0	12	**68**

Table 2. Confusion matrix for maximum accuracy of 70%with FS2 for MLP classifier

This experimentation revealed the need for additional features to improve accuracy. Novel audio features with for expressiveness using articulation in music to capture emotions are proposed for western music. [2] This motivated us to explore the use of music expressive features for emotion perception for Indian music. Traditionally Hindustani classical music uses various music ornamentation forms to convey specific rasa or mood during raga performance. Various ornamentation forms include kan swar, meend, khatka, murkhi,andolan, gamaka, etc [3]. These ornamentations are used in Indian film songs as a natural reflection because of classical music background by various composers and singers. As per the inputs from the performers, kan swar(very small duration note) is used prominently in Indian music for conveying emotional appeal. kan swar is sung or played before or after the main note during the melody. We are experimenting on capturing kan swar from the songs and representing it as a feature dimension. Further, such feature dimensions may be associated with different emotion classes. The proposed hypothesis is by capturing the expressive features of ornamentation, the MER classification accuracy can be improved.

ACKNOWLEDGMENTS

We express our sincere thanks to PICT, CCOEW, SPPU Pune for infrastructural support provided. We thank Dr. H.V. Sahasrabuddhe and Indian music artists for their valuable inputs regarding music emotion.

REFERENCES

[1] Bharata Muni and Manomohan Ghosh. *The Natyasastra*. Manisha Granthalaya.

[2] Renato Panda, Ricardo Manuel Malheiro, and Rui Pedro Paiva. Novel audio features for music emotion recognition. *IEEE Transactions on Affective Computing*, 2018.

[3] Santosh Kumar Pudaruth. A reflection on the aesthetics of indian music, with special reference to hindustani raga-sangita. *SAGE Open*, 6(4):2158244016674512, 2016.