

Unsupervised Approach to Hindi Music Mood Classification

Braja Gopal Patra¹, Dipankar Das², and Sivaji Bandyopadhyay¹

¹Dept. of Computer Science & Engineering, Jadavpur University, Kolkata, India

²Dept. of Computer Science & Engineering, NIT Meghalaya, India
{brajagopal.cse,dipankar.dipnil2005}@gmail.com,
sivaji_cse_ju@yahoo.com

Abstract. We often choose to listen to a song that suits our mood at that instant because an intimate relationship presents between music and human emotions. Thus, the automatic methods are needed to classify music by moods that have gained a lot of momentum in the recent years. It helps in creating library, searching music and other related application. Several studies on Music Information Retrieval (MIR) have also been carried out in recent decades. In the present task, we have built an unsupervised classifier for Hindi music mood classification using different audio related features like rhythm, timber and intensity. The dataset used in our experiment is manually prepared by five annotators and is composed of 250 Hindi music clips of 30 seconds that consist of five mood clusters. The accuracy achieved for music mood classification on the above data is 48%.

Keywords: Hindi Music, Music Mood Classification, MIR, Mood Taxonomy.

1 Introduction

With the rapid evolution of technology, most of the people enhance their life with several technological stuffs. Nowadays, people are enjoying music at leisure time and the overall collection of music increases day by day. However, people are more interested in creating music library which allows them to access songs in accordance with the music moods rather than their title, artists and or genre. Thus, classifying and retrieving music with respect to emotions or mood has become an emerging research area.

Music, also referred as the “language of emotion” can be categorized in terms of its emotional associations [17]. Music perception is highly intertwined with both emotion and the context [6]. The emotional meaning of the music is subjective and thus, it depends upon many factors including culture [7]. Moreover, the mood category of a song varies depending upon several psychological conditions of the Human Beings. Representations of music mood with the psychology remain an active topic for research.

In this paper, we have used a fuzzy c-means classifier for automatic mood classification of Hindi music. As Hindi is the national language of India, Hindi songs are one

of the most popular categories of Indian songs and are present in Hindi cinemas or Bollywood movies. Hindi songs make up 72% of the music sales compared to other language songs in India¹. Mainly, we have concentrated on the collection of Hindi music data annotated with five mood classes². Then, a computational model has been developed to identify the moods of songs using several high and low level audio features. Finally, we have employed fuzzy c-means clustering algorithm and achieved 48% of accuracy on a data set of 250 songs consisting of five mood clusters.

The rest of the paper is organized in the following manner. Section 2 briefly discusses the related work on different languages like English, Indian and Chinese available to date. The dataset and mood taxonomy used in the experiments are described in Section 3. Section 4 describes the list of features for implementing machine learning algorithm. Brief discussion on fuzzy c-means clustering algorithm is described in Section 5. Section 6 presents the experiments with detailed analysis of results. Finally, conclusions are drawn and future directions are presented in Section 7.

2 Related Work

Music classification has received much attention by the researchers in MIR research in the recent years. In the MIR community, Music Information Retrieval Evaluation eXchange³(MIREX) is an annual competition on several important music information retrieval tasks since 2004. The music mood classification task was included into MIREX in the year of 2007. Many tasks were presented related to English music classification such as Genre Classification, Mood classification, Artist Identification, Instrument Identification and Music Annotation etc.

Considerable amount of work has been done on the music mood classification based on audio, lyrics, social tags and all together or in a multi modal approach as described in [6, 16, 17]. Many tasks have been carried out on the English music mood classification such as lyrics [13, 14], audio [7, 18] and both [1, 6]. Some of the works in Chinese music have been conducted based on audio [3] and lyrics [16].

In contrast to other languages, only a few works on mood detection in Indian music has been reported to date and most of the work can be seen on the Carnatic Music [20, 21]. The performance of mood classification on Indian Classical Music was done in Koduri [20] and Hampiholi [21]. However, there are few works available in Hindi Music Mood Classification based on audio [22] and lyrics and can be seen in [19]. Velankar and Sahasrabuddhe [8] prepared data for Hindustani classical music mood classification. They have performed several sessions for classifying the three Indian Ragas into 13 mood classes. To the best of our knowledge, the fuzzy c-means clustering has not been used in mood classification tasks except the work described in [11] where the similar algorithm was used for genre classification of English songs into 10 genres.

¹ http://en.wikipedia.org/wiki/Music_of_India

² The term class and cluster are used interchangeably in this paper.

³ http://www.musicir.org/mirex/wiki/MIREX_HOME

3 Mood Taxonomy and Data Set

One of the issues closely related with mood classification is to identify the appropriate taxonomy for classification. Ekman [9] has defined six basic emotion classes such as *happy, sad, anger, fear, surprise* and *disgust*, but these classes have been proposed for the image emotion classification as we cannot say a piece of music is *disgust*. Thus, in music psychology, our traditional approach is to describe moods using the adjective like *gloomy, pathetic* and *hopeful* etc. However, there is no standard taxonomy available which is acceptable to all the researchers.

Russel [5] proposed the circumplex model of affect based on the two dimensional model. These two dimensions are denoted as “pleasant-unpleasant” and “arousal-sleep”. There are 28 affect words in Russel’s circumplex models and are shown in Figure 1. Later on, Thayer [10] adapted Russel’s model using the two dimensional energy-stress model. Different researchers used their own taxonomy, which are the subsets of Russel’s taxonomy. For example, Katayose et al. [4] used all the adjectives including *Gloomy, Urbane, Pathetic* and *Serious*. Yang et al., [16] used *Contentment, Depression, Exuberance* and *Anxious/Frantic* as mood taxonomy.

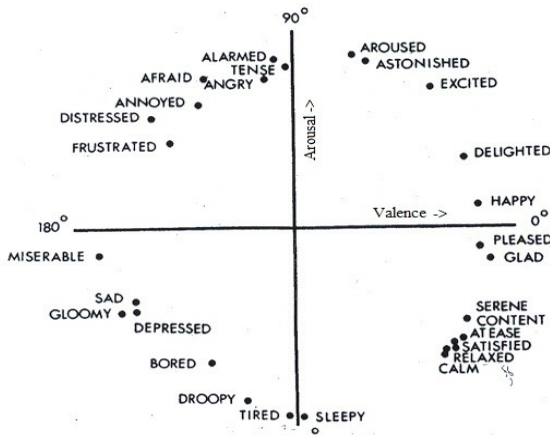


Fig. 1. Russell’s circumplex model of 28 affects words

Table 1. Five mood cluster of proposed mood taxonomy

Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Excited	Delighted	Calm	Sad	Alarmed
Astonished	Happy	Relaxed	Gloomy	Tensed
Aroused	Pleased	Satisfied	Depressed	Angry

Our collected data set includes five clusters of moods according to Theyer’s model [10] and Russel’s Circumplex model [5]. We have also followed the MIREX mood taxonomy [12], which has five mood clusters and each of the clusters has more than

four sub classes. It has been observed that MIREX evaluation forum provides a standard taxonomy for mood classification and many researchers have also used this mood taxonomy [2, 12]. Our mood taxonomy contains five clusters with three sub-classes. The mood taxonomy is formed by clustering similar affect words of Russels' circumplex model. For example, we have kept *calm*, *relaxed* and *satisfied* in one cluster so as to collect similar songs into one group and in this case, the audio features do not vary much. The mood taxonomy used in our experiment is shown in Table 1.

In the present task, a standard data set has been used for the mood classification task. This data has been collected manually and prepared by five human annotators. The songs used in the experiments are collected from Indian Hindi music website⁴. We have faced several problems during the annotation of music. First problem was whether it would be better to ignore the lyrics or not. In Hindi music, we have observed that several songs contain different music as well as different lyrics. For example, a music having high valence consists of the lyric that belongs to the sad mood class. Hu et al. [12] prepared the data based on music only and the lyrics of the song were not considered in their work. So, we also tried to avoid the lyrics of the song as much as possible to build a ground-truth set.

On the other hand, the second problem was the time frame for a song. We have considered only the first 30 seconds of the song so as to prepare our data. In this frame, some lyrics might be present for some of the songs. We have only included the songs that contain lyrics of less than 10 seconds. Finally, we have collected total 250 music tracks out of which 50 tracks were considered from each of the mood clusters. As we have considered only the 30 seconds music from the whole track, it was difficult to identify the track by the annotators. So the inter-annotator agreement was less and was around 72%.

4 Feature Selection

The feature selection always plays an important role in building a good pattern classifier. Thus, we have considered the key features like intensity, timbre and rhythm for our mood classification task. It has been observed that tempo, sound level, spectrum and articulation are highly related to various emotional expressions. Different patterns of the acoustic cues are also associated with different emotional expressions. For example, exuberance is associated with fast tempo, high sound and bright timbre whereas sadness is with slow tempo, low sound and soft timbre.

Rhythm Feature: Rhythm strength, regularity and tempo are closely related with people's moods or responses [3]. For example, generally, it is observed that, in *Exuberance* cluster, the rhythm is usually strong and steady; tempo is fast, whereas in *Depression* cluster, it is usually slow and without any distinct rhythm pattern.

Intensity Feature: Intensity is an essential feature in mood detection and is used in several research works [3, 7]. Intensity of the *Exuberance* cluster is high, and low in

⁴ http://www.songspk.name/bollywood_songs.html

Depression cluster. It is observed that the intensity is approximated in general by considering the root mean square (RMS) values of the corresponding signal.

Timbre Feature: Many existing researchers have shown that mel-frequency cepstral coefficients (MFCCs), so called spectral shapes and spectral contrast are the best features for identifying the mood of music [3, 7, 18]. In this paper, we have used both spectral shape and spectral contrast. Spectral shape includes brightness or centroid, band width, roll off and spectral flux. Spectral contrast features includes sub-band peak, sub-band valley, sub-band contrast.

All the features used in our experiments are listed in Table 2. These features are extracted using jAudio⁵ toolkit. It is a music feature extraction toolkit developed in JAVA platform. The jAudio toolkit is publicly available for research purpose.

Table 2. Feature used in mood classification

Class	Features
Rhythm	Rhythm strength, regularity and tempo
Timbre	MFCCs, Spectral shape, Spectral contrast
Intensity	RMS energy

5 Fuzzy Clustering

We have built an unsupervised classifier to classify the music files into five clusters as stated above in Section 3. We implemented Fuzzy C-means clustering algorithm for the classification purpose. The membership functions of each music vary in between 0 and 1. A detail of the algorithm is given below.

5.1 Fuzzy C-Means Clustering Algorithm

For unsupervised fuzzy clustering, we have chosen the well-known fuzzy c-means clustering algorithm which was already used in music genre classification task [11].

Let there are N data points i.e., $N = \{x_1, x_2, x_3, \dots, x_n\}$ and each data points are represented by p dimensional feature space i.e. $x_k = \{x_{k1}, x_{k2}, x_{k3} \dots x_{kp}\}$.

The main objective of fuzzy c-means clustering algorithm is to classify p dimensional data points N into a set of c fuzzy classes of centroids $v_1, v_2, v_3 \dots v_c$ in same feature space, such that the sum of membership function of any data point x_k , in all classes is 1 [11].

Membership function can be represented by,

$$\sum_{i=1}^c \mu_{ik} = 1, \text{ for } k=1 \text{ to } N. \quad (1)$$

For all c clusters, we can derive the cluster centers v_i for $i=1$ to c by given equation.

⁵ <http://sourceforge.net/projects/jmir/files/>

$$v_i = \frac{\sum_{k=1}^N \mu_{ik}^p x_k}{\sum_{k=1}^N \mu_{ik}} \quad (2)$$

From the above equation, we can see that the cluster center v_i is basically the weighted average of the membership μ_{ik} . To find out the optimal distribution of points of the clusters and optimal placement of centroids, we require an objective function J_p over $\mu = \{\mu_{ik}\}$ and $v = \{v_i\}$, and can be represented as:

$$J_p(\mu, v) = \sum_{i=1}^c \sum_{k=1}^N \mu_{ik}^m \|x_k - v_i\|^2 \quad (3)$$

Where $\| \cdot \|$ is the inner product induced norm in p dimension and Here $m > 1$ is any real number and it influences the membership grade. Then the membership function is updated by the expression given below:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{\|x_k - v_i\|}{\|x_k - v_j\|} \right)^{\frac{2}{m-1}}} \quad (4)$$

In the above experiment, we have used $m=2$, which influence the membership grade. To get desired result, we performed the algorithm iteratively:

- Assigning random values to all v_i and μ_{ik} at initial stage.
- Iteratively recalculate the values for all v_j and then all μ_{ik} according to equation (2) and (4).
- Stop, when the objective function \mathcal{J} changes from the previous iteration less than by a small number δ , a given parameter (here, we used $\delta = 0.01$).

6 Experiments and Evaluation

In order to achieve good results, we require a huge amount of mood annotated music corpus for applying the clustering algorithm. But, to the best of our knowledge, no mood annotated Hindi songs are available to date. Thus, we have developed the dataset by ourselves and it contains 250 songs consisting of five clusters. We have used the fuzzy c -means clustering algorithm described in Section 5 to accomplish our classification experiments based on the features we discussed in Section 4. The features are extracted using the jAudio Feature Extractor.

The accuracies have been calculated and are reported in Table 3. We have considered that a song belongs to *cluster 1* if membership function of *cluster 1* ($\mu_{cluster1}$) is greater than all other membership functions. The confusion matrix of the classification accuracy is given in Table 4. We have achieved the maximum accuracy of 52% in *cluster 3*. It has been observed that the *cluster 1* and *cluster 5* have lowest accuracy and is about 46%. The accuracies of *cluster 2* and *cluster 4* are same and is 48%.

We have observed that some of the instances from each of the clusters have tendency to go towards its neighboring clusters. For example, some songs from *cluster 2* fall under the *cluster 1* as they have similar RMS energy and tempo. The accuracy achieved by the system is quite low as compared to the other existing mood classification systems for

Table 3. Accuracies of each class

Class	Accuracy
Cluster 1	46
Cluster 2	48
Cluster 3	52
Cluster 4	48
Cluster 5	46
Average	48

Table 4. Confusion matrix for the accuracy

		Predictions				
		1	2	3	4	5
T r u e	Clusters	1	2	3	4	5
	1	23	8	4	3	12
	2	11	24	3	4	8
	3	3	10	26	8	3
	4	3	3	13	24	7
5	15	3	5	4	23	

English songs [3, 7], Chinese songs [16], Hindi Songs [19, 22] and Carnatic songs [20, 21]. Later on, inclusion of additional features and the feature engineering may remove such kind of biasness and improve the results.

7 Conclusion and Future Work

In this paper, we have developed the fuzzy c-means classifier for Hindi music mood classification based on the simple audio features namely rhythm, intensity and timbre. We have used our own mood taxonomy described in Section 3 and tried to generalize in accordance to MIREX mood taxonomy and Russel's Circumplex model. We have achieved low accuracy of 48%.

There are several directions for future work. One of them is to improve the fuzzy c-means clustering algorithm using modified objective function as stated in [11]. Incorporate more audio features for enhancing the current results of mood classification. Later on lyrics of the song may be included for multimodal mood classification. Preparing the large audio data and collecting the lyrics of those songs may be considered as the other future direction of research.

References

1. Laurier, C., Grivolla, J., Herrera, P.: Multimodal music mood classification using audio and lyrics. In: Proceedings of the Seventh International Conference on Machine Learning and Applications (ICMLA 2008), pp. 688–693. IEEE (2008)
2. Laurier, C., Sordo, M., Herrera, P.: Mood cloud 2.0: Music mood browsing based on social networks. In: Proceedings of the 10th International Society for Music Information Conference (ISMIR 2009), Kobe, Japan (2009)
3. Liu, D., Lu, L., Zhang, H.J.: Automatic Mood Detection from Acoustic Music Data. In: Proceedings of the International Society for Music Information Retrieval Conference, ISMIR 2003 (2003)
4. Katayose, H., Imai, H., Inokuchi, S.: Sentiment extraction in music. In: Proceedings of the 9th International Conference on Pattern Recognition, pp. 1083–1087. IEEE (1988)
5. Russell, J.A.: A Circumplex Model of Affect. *Journal of Personality and Social Psychology* 39(6), 1161–1178 (1980)
6. Bischoff, K., Firan, C.S., Paiu, R., Nejdil, W., Laurier, C., Sordo, M.: Music Mood and Theme Classification—a Hybrid Approach. In: Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009), pp. 657–662 (2009)

7. Lu, L., Liu, D., Zhang, H.J.: Automatic mood detection and tracking of music audio signals. *IEEE Transactions on Audio, Speech, and Language Processing* 14(1), 5–18 (2006)
8. Velankar, M.R., Sahasrabudde, H.V.: A Pilot Study of Hindustani Music Sentiments. In: *Proceedings of 2nd Workshop on Sentiment Analysis Where AI Meets Psychology (COLING 2012)*, IIT Bombay, Mumbai, India, pp. 91–98 (2012)
9. Ekman, P.: Facial expression and emotion. *American Psychologist* 48(4), 384–392 (1993)
10. Thayer, R.E.: *The Biopsychology of Mood and Arousal*. Oxford University Press, Oxford (1989)
11. Poria, S., Gelbukh, A., Hussain, A., Bandyopadhyay, S., Howard, N.: Music Genre Classification: A Semi-supervised Approach. In: Carrasco-Ochoa, J.A., Martínez-Trinidad, J.F., Rodríguez, J.S., di Baja, G.S. (eds.) *MCPR 2012. LNCS*, vol. 7914, pp. 254–263. Springer, Heidelberg (2013)
12. Hu, X., Downie, S.J., Laurier, C., Bay, M., Ehmann, A.F.: The 2007 MIREX Audio Mood Classification Task: Lessons Learned. In: *Proceedings of the 9th International Society for Music Information Retrieval Conference (ISMIR 2008)*, pp. 462–467 (2008)
13. Hu, X., Downie, S.J., Ehmann, A.F.: Lyric text mining in music mood classification. In: *Proceedings of 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, pp. 411–416 (2009)
14. Hu, Y., Chen, X., Yang, D.: Lyric-based Song Emotion Detection with Affective Lexicon and Fuzzy Clustering Method. In: *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, pp. 123–128 (2009)
15. Yang, Y.H., Liu, C.C., Chen, H.H.: Music emotion classification: a fuzzy approach. In: *Proceedings of the 14th Annual ACM International Conference on Multimedia*, pp. 81–84. ACM (2006)
16. Yang, Y.-H., Lin, Y.-C., Cheng, H.-T., Liao, I.-B., Ho, Y.-C., Chen, H.H.: Toward multimodal music emotion classification. In: Huang, Y.-M.R., Xu, C., Cheng, K.-S., Yang, J.-F.K., Swamy, M.N.S., Li, S., Ding, J.-W. (eds.) *PCM 2008. LNCS*, vol. 5353, pp. 70–79. Springer, Heidelberg (2008)
17. Kim, Y.E., Schmidt, E.M., Migneco, R., Morton, B.G., Richardson, P., Scott, J., Speck, J.A., Turnbull, D.: Music emotion recognition: A state of the art review. In: *Proceedings of 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, pp. 255–266 (2010)
18. Fu, Z., Lu, G., Ting, K.M., Zhang, Z.: A survey of audio-based music classification and annotation. *IEEE Transactions on Multimedia* 13(2), 303–319 (2011)
19. Ujlambkar, A.M., Attar, V.Z.: Mood classification of Indian popular music. In: *Proceedings of the CUBE International Information Technology Conference*, pp. 278–283. ACM (2012)
20. Koduri, G.K., Indurkha, B.: A Behavioral Study of Emotions in South Indian Classical Music and its Implications in Music Recommendation Systems. In: *Proceedings of the 2010 ACM Workshop on Social, Adaptive and Personalized Multimedia Interaction and Access*, pp. 55–60. ACM (2010)
21. Hampiholi, V.: A method for Music Classification based on Perceived Mood Detection for Indian Bollywood Music. *World Academy of Science, Engineering and Technology* 72, 507–514 (2012)
22. Patra, B.G., Das, D., Bandyopadhyay, S.: Automatic Music Mood Classification of Hindi Songs. In: *Proceedings of 3rd Workshop on Sentiment Analysis where AI meets Psychology (IJCNLP 2013)*, Nagoya, Japan, pp. 24–28 (2013)