

A method for Music Classification based on Perceived Mood Detection for Indian Bollywood Music

Vallabha Hampiholi

Abstract—A lot of research has been done in the past decade in the field of audio content analysis for extracting various information from audio signal. One such significant information is the "perceived mood" or the "emotions" related to a music or audio clip. This information is extremely useful in applications like creating or adapting the play-list based on the mood of the listener. This information could also be helpful in better classification of the music database. In this paper we have presented a method to classify music not just based on the meta-data of the audio clip but also include the "mood" factor to help improve the music classification. We propose an automated and efficient way of classifying music samples based on the mood detection from the audio data. We in particular try to classify the music based on mood for Indian bollywood music. The proposed method tries to address the following problem statement: Genre information (usually part of the audio meta-data) alone does not help in better music classification. For example the acoustic version of the song "nothing else matters by Metallica" can be classified as melody music and thereby a person in relaxing or chill out mood might want to listen to this track. But more often than not this track is associated with metal / heavy rock genre and if a listener classified his play-list based on the genre information alone for his current mood, the user shall miss out on listening to this track. Currently methods exist to detect mood in western or similar kind of music. Our paper tries to solve the issue for Indian bollywood music from an Indian cultural context.

Keywords—Mood, music classification, music genre, rhythm, music analysis.

I. INTRODUCTION

AFTER SILENCE, that which comes nearest to expressing the inexpressible is music. So said one of the greatest novelist of our times Aldous Huxley about music. Music is the ordering of tones or sounds to produce compositions with unity and continuity. Musical compositions utilize the five elements of music (*rhythm, melody, pitch, harmony and interval*), which play a significant role in human physiological and psychological functions, thus creating alterations in mood [1]. People listen to music in their leisure time and what kind of music people listen to is governed by what "mood" they are in. For example a person in angry mood is more likely to listen to music of genre heavy metal and hard rock. In today's world there are a variety of music sources like, portable music players, USB based storage devices, SD Cards, mobile phones, internet (cloud), radio (AM/FM/DAB/SDARS) available to music listeners. With a variety of music sources, music control (organization and management) plays a crucial role in music search and play-list creation for music listeners.

Vallabha Hampiholi is with the automotive division of Harman International, Bengaluru, India e-mail: vallabha.hampiholi@harman.com

Although some information like album name, artist name, genre, year of album release are present in the meta-data of the music clip, but their significance is limited when it comes to music related applications like creation of play-list based on the mood of the listener. To address such scenarios we need more information than what is available in the meta-data like beat, mood and tempo [2].

A lot of research has been done in the area of beat, tempo and genre detection and classification using various features of the audio content. Ellis et al. [3] presented a beat-tracking system for identifying 'cover' songs, and Scheirer [4] proposed a band-pass filtering and parallel comb filtering based beat tracking system. Peeters in [5] proposed a method to define the rhythm of audio signal based on the spectral and temporal periodicity representations. Researchers in [6] present a technique to classify audio genre based on clustering sequences of bar-long percussive and bass-line patterns. Mancini et al. [7] explore a method to extract the various expressions inherent in a musical piece. Lie et al in [8] proposed a method to automatically detect the mood from music data using psychological theories in western cultures. Various related works on the audio content analysis can be found in the proceedings of ISMIR [9], a conference for music information retrieval.

As seen from the various research topics presented above, few works have focussed on accurately classifying the music based on the audio content information that directly relates to the mood of the listener especially for Indian cultural context. In this paper, we try to interpret the mood in a musical piece based on the empirical data acquired through listening tests. Adding this new parameter for music classification, it would help the listener to listen to variety of artists whose music exhibit the moods similar to that of the listener rather than just using the traditional meta-data information as done previously.

A. Motivation

A mood is a relatively long lasting emotional state. Moods differ from emotions in that they are less specific, less intense, and less likely to be triggered by a particular stimulus or event [10]. Picard in [11] talks about affective computing which is about building intelligent systems that can recognize, interpret, process, and simulate human feeling or emotion. The motivation behind Picard's research is the ability to simulate empathy.

Researchers in [12] show that humans are not only able to recognize the emotional intentions used by musicians but

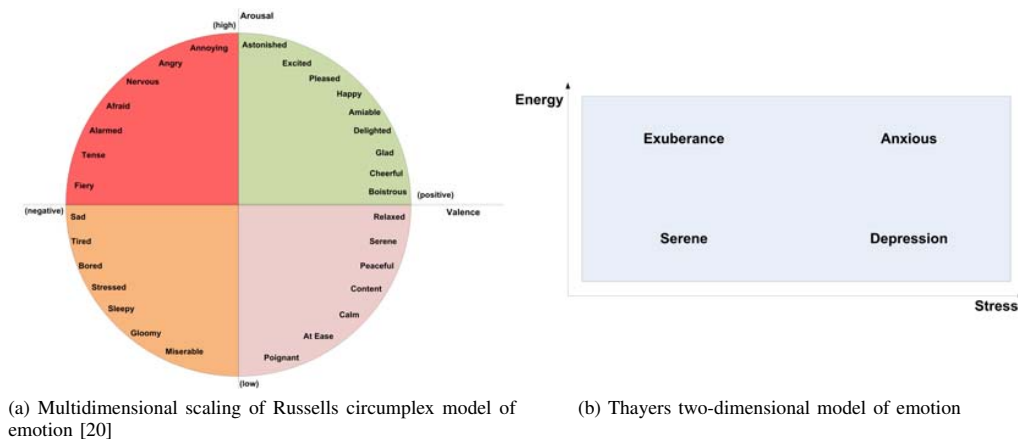


Fig. 1: Illustration of Russell's circumplex and Thayer's 2-dimensional model of mood.

also feel them. When we listen to various music we normally tend to experience a change in blood pressure, heart beat rate etc. In this paper we try to present a method to classify music based on the mood of the user. To illustrate it, we have developed a media player that can classify music (mp3 songs) based on popular meta data information like genre and the perceived mood in the song. Our experiments show that the mood information helps in better classification of the songs than using the traditional information like genre.

Possible applications of the proposed music classification based on mood are:

- 1) Intelligent media players (portable) which could take the user mood as additional parameter to generate / modify existing play list.
- 2) Intelligent systems (cloud based) to recommend songs based on the user current mood.
- 3) Cloud based media players which could adapt the play list for the user based on the user mood at different locations (office, home, recreational places etc.)
- 4) Intelligent media players (portable or cloud based) which could learn about the mood the user is in based on the songs he/she has listened to and recommend the next song accordingly.

II. CHALLENGES

Some previous works like [5] [13] have addressed the methods for music classification. However the work by [13] is focussed on the music label categorization based on different categories (*Style, Genre, Musical set-up, Main instruments, Variant, Dynamics, Tempo, Era/Epoch, Metric, Country, Situation, Mood, Character, Language, Rhythm, and Popularity*) based on a corrective approach. The research assumes that the category information is already available in the meta-data and the algorithm tries to correct the information based on the acoustical data analysis. Also the work in [5] is related to classifying music based on the spectral and temporal periodicity in the audio data alone.

Our work is based on extracting or estimating the inherent mood in the musical piece by extracting various parameters

like *beat, tempo, rhythm, timbre, pitch, tone, sound level, genre and vibrato* and tagging it to the song. We then use this information to classify the songs in a media player. We have written a MATLAB based media player which can classify a user play-list using the genre or mood information. And as per our experiments on a song database containing 250 mp3 tracks we found that the play-list generated using the mood information listed songs that aligned with the user mood than using the standard genre information alone. The work presented in the paper tries to map the expression in the music to the mood of the user. The mapping has been validated through the MATLAB based mp3 player which modifies a given play-list based on the user mood information.

A. Mood Detection - Challenges

To build an efficient model for mood detection using the acoustical data there are many challenges that need to be considered. The following challenges are significant

- **Mood Perception** Properties of music, might provide universal cues to emotional meaning. But it is difficult to assess this hypothesis, because most research on the topic has involved asking listeners to judge the emotional meaning of music which is subjective and depends on factors like culture, education and individual personality [15]. Hence, for the same musical piece, different individuals might have different perceptions. Further researchers in [16] show that the variance in perception of mood inherent in a musical piece within a given cultural context could be minimal. Therefore, it is possible to build a mood detection system in a certain context. In this paper, our system is based on Indian cultural context.

- **Mood Classification**

There is always a debate on what kind of emotion a musical piece can express and whether it is perceived by us. Researchers like in [14] use various adjectives to describe mood. However research work like [10] also provide the basis for mood classification. We adopt the Thayer's model [10] for mood classification.

TABLE I: Acoustic Cues in Musical Performance [16] [17].

Audio Feature	Exuberance	Anxious	Serene	Depression
Rhythm(Tempo)	Regular(Fast)	Regular(Fast)	Irregular(Slow)	Irregular(Slow)
Sound Level/Energy	High	High	Very Low	Low
Spectrum	Medium HF Energy	High HF Energy	Little HF Energy	Little HF Energy
Timbre	Bright	Sharp	Dull	Dull
Articulation	Staccato	Legato	Staccato	Legato

• Acoustic Cues

In our work we have dealt with decoded audio data from which it is difficult to interpret mood directly. Many a work have been proposed before in extracting significant features like mel-frequency cepstral coefficients (MFCC), signal energy, zero crossing rate (ZCR) etc., from the acoustic signal. Therefore, to build an efficient model for mood detection in music, we need to extract acoustic features to represent the basis of various moods. In our method, intensity (energy), timbre (ZCR, Brightness) and rhythm (tempo) features are extracted, from the acoustic signal to map it to a particular classification of mood. We have used the MIRtoolbox [18] to extract the aforementioned features from the musical piece.

Our approach of mood detection and tagging the mood information to the musical piece, is formed by considering the above mentioned challenges and solutions.

B. Mood Classification

As has been discussed before, mood in contrary to emotions, is not directed at a specific object. "When one has the emotion of sadness, one is normally sad about a specific state of affairs, such as an irrevocable personal loss. In contrast, if one is in a sad mood, one is not sad about anything specific, but sad "about nothing (in particular)" [19]. In this paper to distinguish between mood and emotions, it is assumed that an emotion is a subset of the mood. A mood can comprise of several emotions.

One of the most challenging factor of mood detection in a musical piece is mood classification. There has been lot of study done when it comes to emotion classification [11]. Mood is a very subjective notion, and hence there is no standard mood classification system that is accepted by all. Xiao et al [14] propose set of 5 clusters each having various adjectives to describe the mood. The mood clusters effectively reduce the diverse mood space into a tangible set of categories.

Russell [20] proposed a circumplex model of affect based on two bipolar dimensions. The two dimensions are called pleasant-unpleasant and arousal-sleep. Thus, each affect word could be defined as some combination of the pleasure and arousal components. According to Thayer [10] adapted Russells model to music using two dimensions: energy and stress as shown in Figure 1b. In Thayer's model, the energy (vertical dimension) corresponds to the arousal, while stress (horizontal dimension) corresponds to pleasure in Russells model Figure 1a.

Based on Thayer's model of energy(arousal) vs stress(valence) music mood can be divided into four clusters namely: *Contentment, Depression, Exuberance and Anxious* as shown in Figure 1b. These four clusters are explicit and easily discriminable. Hence this model of mood classification is applied in our mood detection algorithm. To be consistent, throughout the remaining paper we shall continue to use the adjectives of Thayer's model.

From Figures 1a and 1b we could map the adjectives of Thayer model to the unique quadrants of Russell's model. Anxious mood maps to the emotions like nervousness, afraid etc and hence could be mapped to the quadrant coloured red. Similarly exuberance maps to the green coloured quadrant, depression to orange quadrant and serene to violet quadrant.

III. LISTENING TESTS - FOR INDIAN CULTURAL CONTEXT BASED MODELLING

As stated earlier, our mood detection system is based on the Indian cultural context. In order to tune our algorithm, we conducted some listening tests. The listening test initially included 145 bollywood song clips each of duration 1 minute. The test audience included 30 people (22, Male listeners and 8, Female listeners with an average age of 28 years). All the listeners were from Indian cultural context. Listeners were asked to identify the emotions inherent in the tracks and rate them on a scale of 0 - 9. Ratings were rounded off at mid-range value of 4 i.e, if a user rated a musical piece as 3 on anxiety and 7 on exuberance, the musical piece's mood was classified as exuberance. Also for each song the rating were arrived at using the average values as assigned by the listeners. If a song were perceived to have more than one mood meaning the listeners had different opinion, the song was removed from our training database. The total number of songs after the short listing was 122. Table II shows the classification of the music database into different moods based on the listening tests.

TABLE II: Classification Based On Listening Test Results

Sl.No	Mood	Number of Songs	%
1	Exuberance	43	35.24
2	Anxious	20	16.39
3	Serene	28	22.95
4	Depression	31	25.40

IV. AUDIO FEATURE EXTRACTION

Acoustic audio data can digitally be represented in either frequency domain (spectral analysis) or time domain (temporal analysis). In the frequency domain, spectral descriptors are often computed from the Fourier Transform (FT). Many acoustic features can be derived from the FT [18]:

- Basic statistics of the spectrum gives some timbral characteristics (such as spectral centroid, roll-off, brightness, flatness, etc.).
- The temporal derivative of spectrum gives the spectral flux.
- An estimation of roughness, or sensory dissonance, can be assessed by adding the beating provoked by each couple of energy peaks in the spectrum
- A conversion of the spectrum in a Mel-scale can lead to the computation of Mel-Frequency Cepstral Coefficients (MFCC)
- Tonality can also be estimated

One of the simplest features, zero-crossing rate (ZCR), is based on a simple description of the audio waveform itself: it counts the number of sign changes of the waveform. Signal energy is computed using root mean square, or RMS [21].

A. MIRtoolbox for Audio Feature Extraction

MIRtoolbox has been developed within the context of a European Project called "Tuning the Brain for Music", funded by the NEST (New and Emerging Science and Technology) program of the European Commission. MIRtoolbox offers an integrated set of functions written in Matlab, dedicated to the extraction from audio files of musical features such as tonality, rhythm, structures, etc. The objective is to offer an overview of computational approaches in the area of Music Information Retrieval. Various acoustic information like Zero-crossing rate, RMS energy, MFCC, pitch, tempo etc., can be extracted from audio data using MIRtoolbox [18].

In our system we extract the following audio features from the musical data:

- *Audio Intensity/Energy*: From Thayer's model depicted in Figure 1b it is easy to see why audio signal intensity is very important when it comes to mood detection. Musical pieces with high intensity relate to *Exuberance* and *Anxious* whereas with low intensity relate to *Serene* and *Depression*. The above observations about signal intensity has also been noted in [16] as shown in Table I.
- *Timbre*: Timbre or the tone color is the quality of a musical note that distinguishes different styles of music. Timbre makes a particular musical piece sound different from another, even if they have the same pitch and loudness. Its primary component parts are the dynamic envelope, spectrum and spectral envelope [22]. Timbre plays an important role in human perception of music. Tones with many higher harmonics is related to *Anxiousness* and *Exuberance*, whereas with few, low harmonics can be associated with *Serenity* and *Depression* [17].
- *Rhythm*: Human mood response is dictated by the tempo and rhythm periodicity present in a musical piece [16]. Important thing of note is that in a given musical piece

there is no simple relationship between its timbre and rhythm. There are pieces and styles of music which are texturally and timbrally complex, but have straightforward, perceptually simple rhythms; and there also exist musics which deal in less complex textures but are more difficult to rhythmically understand and describe [4]. Regular rhythm with fast tempo may be perceived as expressing exuberance, while irregular rhythm with slow tempo conveys anxiousness [17].

An indicative list of acoustic features that we extract from a musical piece to estimate the conveyed mood is shown in Table III.

TABLE III: Acoustic Features And Their Definitions [23]

Intensity Features	
Energy(RMS)	Represents the power of audio signal.
Low Energy Frames	Number of frames with lower energy than the threshold value indicates extent of quietness in a musical piece.
Timbre Features	
Brightness	Indication of the amount of high-frequency content in a sound, using a measure such as the spectral centroid.
Bandwidth	Indicates the number of instruments used in a musical piece.
Roll-Off	Indication of expression of darkness. Higher low frequency energy in a musical piece expresses sadness and depression. On the contrary brighter and cheerful music is characterized by high high frequency energy. One way to estimate the amount of high frequency in a signal is finding the frequency such that a certain fraction of the total energy is contained below that frequency.
Zero Cross	Indicates the level of noise in the signal. Measured by counting the number of times the signal crosses the X-axis (or, in other words, changes sign).
Rhythm Features	
Fluctuation	Indicates the rhythmic periodicities.
Tempo	Indicates a measure of acoustic self-similarity as a function of time lag. Music exhibiting exuberance and anxiousness tend to have faster tempo compared to musical compositions that exhibit serene and depression moods.

V. PROPOSED METHOD

Based on Thayer's model, a framework for mood detection is proposed. Many a methods have been proposed on Mood detection in music [8] [24]. However these methods have been proven to work well with western classical music or western music in general. We propose a method to detect mood in bollywood (Indian) music as the methods [8] [24] will not work effectively as there is a huge difference between the two contexts culturally and demographically.

A. Approach for Mood Detection

The main aim of this research was to determine the mood in the musical piece. The proposed mood detection technique is based on Thayer's model. It has been shown by researchers

TABLE IV: Musical Features - Thayer Model Mapping

Mood	Intensity Features		Rhythm Features	Timbre Features		
	Mean Energy	Mean Low Energy	Mean Tempo	Mean ZCR	Mean Brightness	Mean Roll-Off (95%)
Exuberance	0.9237	0.4795	135	1422	0.4883	10377
Anxious	0.9468	0.4515	124	1360	0.4644	10287
Serene	0.7284	0.5264	125	1058	0.3822	08508
Depression	0.8467	0.5484	132	1315	0.4038	08661

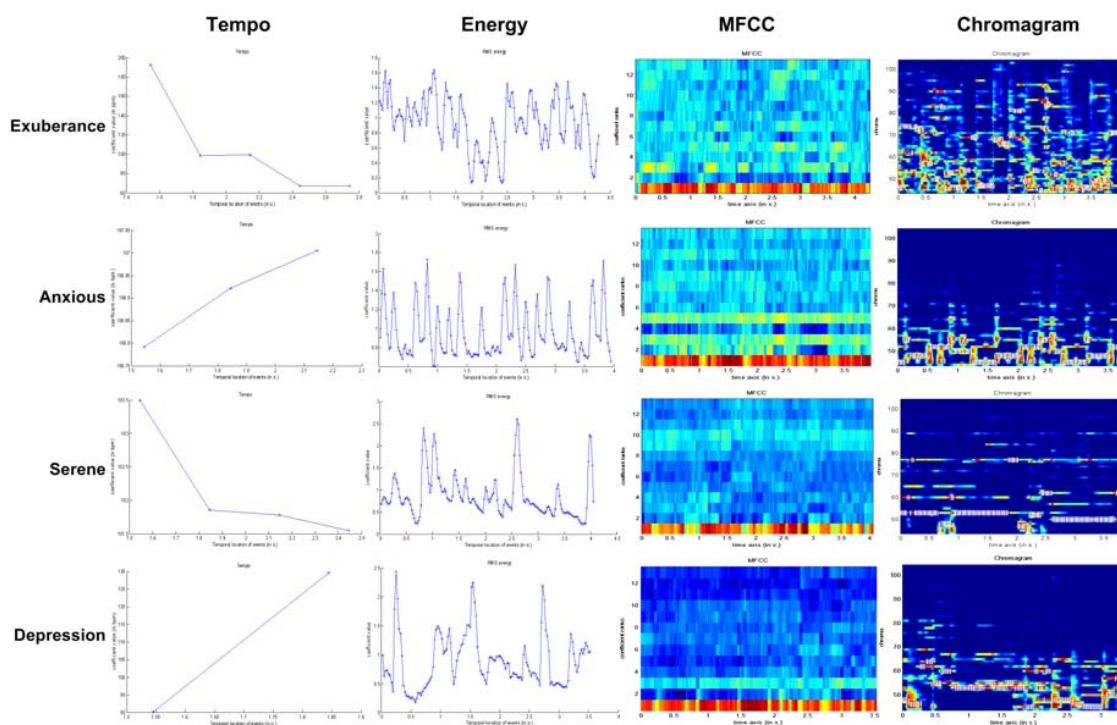


Fig. 2: Graphical representation of sample songs mapping to Thayer’s Model and their respective audio features.

that energy of audio signal is more computationally measurable than the valence(stress) factor. Our approach for mood detection is illustrated in Figure 3a. The first step was to map sample music clips into known Thayer’s mood model through series of listening tests. Machine learning is then used to construct a computational model using the listening test result and the measurable audio features of the musical pieces. As stated earlier the following attributes of the musical pieces are analysed, mood, energy, tempo, ZCR and brightness. The mood attribute is extracted via the listening tests, and the rest of the numerical features are extracted using the mirtoolbox [18] in Matlab. The computational model will be presented in the following Section of the paper.

B. Machine Learning for Mood Detection

As sated in the previous section our proposed system of mood detection is based on machine learning technique. We use the results of the listening tests shown in Table II for machine learning. Machine learning assumes that any hypothesis

found to approximate the target function well over a sufficiently large set of training examples will also approximate the target function well over other unobserved examples. So a system can estimate a mood class in an unobserved sample by comparing it against the observed trained samples.

TABLE V: Confusion Matrix - Generated using C4.5 classifier with 122 instances of which 66% was used to train and remainder for testing.

a	b	c	d	Classified as
10	1	0	3	a = exuberance
2	3	0	1	b = anxiety
0	0	10	1	c = serene
2	0	0	8	d = depression

The classification of a song incorporates the features mentioned in Table IV which are extracted from the audio clip. In order to reduce the computational complexity we consider the

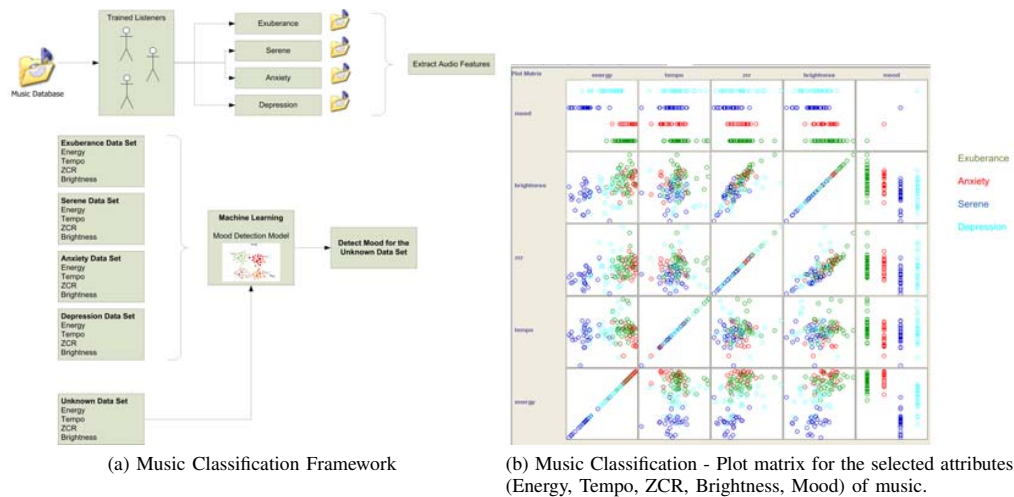


Fig. 3: Illustration of our music classification model and the decision tree to detect the mood in an unknown sample.

Energy, Tempo, ZCR and Brightness attributes of the musical pieces. We have used WEKA [25] tool to generate our music classification model.

Table V shows confusion matrix generated using the C4.5 [26] decision tree learner. As can be seen from the matrix, the *True Positive* rate for each of the mood exuberance, anxiety, serene and depression is .714, .5, .909 and .8 respectively.

The classification model is based on the system generated pruned tree. We perform experiments on untrained data and the results are presented in the following section.

VI. RESULTS

Based on the system computed in the above section we perform a few experiments on the untrained data. The result summary is presented in Table VI. The untrained musical samples are first classified with listening tests, the results of which are presented in column 4 of Table VI. The audience for listening test included 4 males and 3 females with an average age of 30 years. All the audience members were from Indian cultural background. Then each of musical clip was classified using the classification system derived in the above section. The results of the classification system is presented in column 3 and the result of the system based classification Vs listening test is presented in column 5 of Table VI. The untrained data is a mix of Indian bollywood and western music.

As can be inferred from the results, the classification system works particularly well with the Indian bollywood music whereas its performance is not desirable when it comes to western music. The reasoning for this can be summarized as:

- 1) The classification system was based on the listening tests conducted with Indian cultural context.
- 2) The audience to classify the western music were from Indian cultural background. This means that the way they interpreted western music could be entirely different from the way it is perceived in western cultures. For example the song Walk by Pantera is classified as a song exhibiting anxiousness or being frantic. However

the audience in this particular test have classified the song belonging to mood group depression.

- 3) From the confusion matrix in Table V we see that the True Positive rate for the anxiety mood is pretty low at .5. This observation is exhibited in the tests (row 6 and 7) of our experimental results shown in Table VI.
- 4) From the listening test it was also observed that a few songs could not fit into the moods as classified by Thayer's model. This means that for future we need to have more classifications with regards to summarizing human mood efficiently.

From Table VI we observe that the success rate of detecting the mood accurately for Indian bollywood music is 60%. We also observe that the success rate (40%) falls when detecting mood in western music. The success rate for the western music is expected to be lower as our music classification framework was entirely based on the Indian cultural context and the audience classifying the western music through listening tests were all from Indian cultural background thus leading to the observed results which are in accordance with previous studies as shown in [15].

During our experiments and listening tests we have observed that sometimes a musical piece has mixture of moods. It has also been observed that sometimes it's difficult to classify the music based on limited classification groups of Thayer's mood model.

VII. CONCLUSION

We have presented a method to detect mood in Indian bollywood music based on Thayer's mood model comprising four mood types, *Exuberance*, *Anxiety*, *Serene*, and *Depression*. Audio features related to Intensity (Energy), Timbre (Brightness and ZCR), and Rhythm (Tempo) are extracted from the musical data. Music classification model based on mood is arrived at using machine learning technique.

As can be seen from our experimental results shown in Table VI there is a need for further improvement for the proposed method. More audio features could help improve the accuracy

TABLE VI: Mood Classification Experimental Results - System Vs Listening Test.

Sl.No	Track Title	System Mood Classification	Listening Test Result	Test
1	Sooraj Ki Baahon Mein (Zindagi Na Milegi Dobara)	Exuberance	Exuberance	PASS
2	Comfortably Numb (Pink Floyd)	Depression	Depression	PASS
3	Tu Hi Meri Shab Hai (Gangster)	Exuberance	Exuberance	PASS
4	Kaisi Hai Yeh Rut (Dil Chahta Hai)	Serene	Serene	PASS
5	Tanhayee (Dil Chahta Hai)	Anxiety	Depression	FAIL
6	Aa Zara (Murder 2)	Exuberance	Axiety	FAIL
7	Stairway To Heaven (Led Zeppelin)	Serene	Exuberance	FAIL
8	The Ketchup Song (Las Ketchup)	Exuberance	Exuberance	PASS
9	Walk (Pantera)	Anxiety	Depression	FAIL
10	I Dreamed Of A Dream (Susan Boyle)	Anxiety	Depression	FAIL

of the system. We also observe that the data set used to build our classification model could be increased further to improve the accuracy of the classification system.

VIII. FUTURE WORK

We plan to extend the framework to recognize moods related to variety of Indian music like Indian classical music (Carnatic, Hindustani), Ghazals, Qawwali along with bollywood music. This will mean tedious task of collecting variety of musical pieces and having qualified listeners who can interpret the meanings in these kind of compositions accurately.

REFERENCES

- [1] Carolyn J. Murrock, "Music and Mood," in *Psychology of Moods 2005*
- [2] D. Huron, "Perceptual and cognitive applications in music information retrieval," in *Proc. Int. Symp. Music Information Retrieval (ISMIR)*, 2000.
- [3] Ellis, D.P.W.; Poliner, G.E.; , "Identifying 'Cover Songs' with Chroma Features and Dynamic Programming Beat Tracking," *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on* , vol.4, no., pp.IV-1429-IV-1432, 15-20 April 2007
- [4] E. Scheirer, "Tempo and beat analysis of acoustic musical signals," *J. Acoust. Soc. Amer.*, vol. 103, no. 1, pp. 588601, 1998.
- [5] Peeters, G.; , "Spectral and Temporal Periodicity Representations of Rhythm for the Automatic Classification of Music Audio Signal," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol.19, no.5, pp.1242-1252, July 2011
- [6] Tsunoo, E.; Tzanetakis, G.; Ono, N.; Sagayama, S.; , "Beyond Timbral Statistics: Improving Music Classification Using Percussive Patterns and Bass Lines," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol.19, no.4, pp.1003-1014, May 2011
- [7] Mancini, M.; Bresin, R.; Pelachaud, C.; , "A Virtual Head Driven by Music Expressivity," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol.15, no.6, pp.1833-1841, Aug. 2007
- [8] Lie Lu; Liu, D.; Hong-Jiang Zhang; , "Automatic mood detection and tracking of music audio signals," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol.14, no.1, pp. 5- 18, Jan. 2006
- [9] *Proc. ISMIR: Int. Symp. Music Information Retrieval*, [Online]. <http://www.ismir.net/>.
- [10] Thayer, Robert E. (1998). "The Biopsychology of Mood and Arousal". New York, NY: Oxford University Press
- [11] Picard, Rosalind. "Affective Computing" MIT Technical Report #321, 1995
- [12] P. N. Juslin; P. Laukka, "Communication of emotions in vocal expression and music performance: Different channels, same code?," *Psychol. Bull.*, vol. 129, no. 5, pp. 770814, 2003.
- [13] Pachet, F.; Roy, P.; , "Improving Multilabel Analysis of Music Titles: A Large-Scale Validation of the Correction Approach," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol.17, no.2, pp.335-343, Feb. 2009
- [14] Hu, Xiao.; J. Stephen Downie, "Exploring mood metadata: Relationships with genre, artist and usage metadata," *International Conference on Music Information Retrieval (ISMIR 2007)*, Vienna, September 23-27, 2007
- [15] L.-L. Balkwill; W. F. Thompson; R. Matsunag, "Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners," *Japanese Psychological Research*, Volume 46, No. 4, 337349, 2004
- [16] P. N. Juslin, "Cue utilization in communication of emotion in music performance: relating performance to perception," *J. Exper. Psychol.: Human Percept. Perf.*, vol. 16, no. 6, pp. 17971813, 2000.
- [17] Patrik N. Juslin, John A. Sloboda, "Handbook of music and emotion: theory, research, applications"
- [18] Olivier Lartillot; Petri Toivainen, "A Matlab Toolbox For Musical Feature Extraction From Audio.," in *Proc. of the 10th Int. Conference on Digital Audio Effects (DAFx-07)*, Bordeaux, France, September 10-15, 2007
- [19] Siemer M. (2005). "Moods as multiple-object directed and as objectless affective states: An examination of the dispositional theory of moods." *Cognition & Emotion: January 2008*. Vol 22, Iss. 1; p. 815-845.
- [20] J. A. Russell, "A circumplex model of affect," *J. Personality Social Psychology*, vol. 39, pp. 11611178, 1980.
- [21] G. Tzanetakis and P. Cook, "Multifeature audio segmentation for browsing and annotation," in *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 1999.
- [22] William Moylan, "The art of recording: understanding and crafting the mix"
- [23] Rossitza Setchi, "Knowledge-Based And Intelligent Information And Engineering Systems:," *14th International Conference, KES 2010*, Cardiff, UK, September 8-10, 2010
- [24] Eerola, T., Lartillot, O., and Toivainen, P. "Prediction of multidimensional emotional ratings in music from audio using multivariate regression models." In *Proceedings of 10th International Conference on Music Information Retrieval/ISMIR 2009*, pages 621-626.
- [25] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten (2009); "The WEKA Data Mining Software: An Update"; SIGKDD Explorations, Volume 11, Issue 1.
- [26] Ross Quinlan (1993). C4.5: "Programs for Machine Learning, Morgan Kaufmann Publishers", San Mateo, CA.
- [27] Meng, A.; Ahrendt, P.; Larsen, J.; Hansen, L.K.; , "Temporal Feature Integration for Music Genre Classification," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol.15, no.5, pp.1654-1664, July 2007
- [28] Hanjalic, A.; , "Extracting moods from pictures and sounds: towards truly personalized TV," *Signal Processing Magazine, IEEE*, vol.23, no.2, pp.90-100, March 2006
- [29] Changsheng Xu; Maddage, N.C.; Xi Shao; , "Automatic music classification and summarization," *Speech and Audio Processing, IEEE Transactions on*, vol.13, no.3, pp. 441- 450, May 2005



Vallabha Hampiholi (M'08) received his Bachelor's degree in Electronics and communication from Kuvempu University, India in 2000 and Master's degree in Electronic Systems from La Trobe University, Melbourne in 2005. Since 2010, he has been with Automotive division of Harman International, Bengaluru working in areas related to audio signal routing, control and processing.