Linking and Familiarity Rating Method Classifies the Music, Video Assessment Responses of EEG-Signal

B. Senthil Kumar
Research Scholar,
Department of Computer Science of CS,
School of Science & Technology,
Dravidian University,
Kuppam-517425, A.P.
India.

S. Santhosh Baboo
PG & Research Department of Computer Application
DG Vaishnav College,
Arumbakkam, Chennai,
India.

ABSTRACT
The most renowned strategy utilized for perusing mind movement is electroencephalography (EEG). Electroencephalography is the neurophysiologic estimation of the electrical action of the cerebrum by recording from anodes put on the scalp, or in the exceptional cases on the cortex. The ensuing follows are known as an electroencephalogram (EEG) and speak to alleged brainwaves. This system is picking up prevalence as it is a non-intrusive interface and is giving a methodology to controlling machines through contemplations. The proposed linking and familiarity rating method classifies the music, video assessment responses of EEG-Signal. The metrics namely true positive, true negative, false positive, false negative, sensitivity, specificity and classification accuracy are chosen for evaluating the performance of the proposed classifier. The simulation result shows that the proposed classifier achieves 95.4 % accuracy which is better than other methods.

Keywords
Usability, Electroencephalography, EEG, Familiarity rating, Bio feedback, User experience, Music video assessment, responsive signals, BCI, Brain Compute Interface.

1. INTRODUCTION
The area of designing brain-computer interfaces (BCI) is considered challenging and most desired, as it has great potential for physically challenged persons with severe communication and control problems [Adams et al., 2008]. The BCI provide tools that connect human brain directly to a computer device and offers an environment that greatly differs from the traditional way of communication. Generally electroencephalography EEG method is used for reading brain activity. Electroencephalography is the neurophysiologic measurement of the electrical activity of the brain by recording from electrodes placed on the scalp, or in the special cases on the cortex. The resulting traces are known as an electroencephalogram (EEG) and represent so-called brainwaves. This method is gaining popularity as it is a non-invasive interface and is providing an approach for controlling computers through thoughts.

Listening to music is one of the most common forms of interaction between users and multimedia devices (Zajonc.,1968, Borstein.,1989). Often, users rate and file songs according to their personal taste or even receive suggestions for potentially favorite musical content. The incorporation of the latter aspect of musical experience, i.e., music appreciation, in bio-inspired systems, and especially brain computer interfaces (BCIs), fosters the establishment of a direct pathway between brain processes and multimedia interfaces, which would prospectively assist the integration of a more holistic experience of music in the everyday life of patients suffering from cognitive or motor impairments.

The normal way to accept input is through keyboards and mouse, which require operations made by a user. Communication and feedback is usually achieved through verbal (audio) or visual elements like links and pictures. Usability analysis is an area that is directed at making this communication comfortable. Persons with physical inability find it more difficult to access the various parts of a screen and, because of their inability, find it very difficult to provide a feedback on their using experience. A solution to this is to use the concept of Brain- Computer Interface devices to connect computers with severely disabled individuals to develop effective and efficient communication interface. Thus, the primary objective of the research work is to develop a BCI, which can be used by motion disabled people (e.g. amyotrophic lateral sclerosis, brainstem stroke, cerebral palsy and spinal cord injury) and study their user experience for usability analysis.

The objective of this paper is to take a bio-feedback mechanism by the way of non-invasive EEG device and to perform binary classification task from the received EEG signal as calm and excited status along with evaluating the performance using performance metrics namely true positive, true negative, false positive, false negative, sensitivity, specificity and classification accuracy.

2. LIRERATURE REVIEW
In order for such BCI systems to be effective, parameters that pertain to the formation of music preference should be taken into consideration during their design. Music preference constitutes a distinctive characteristic of our personality that is influenced by numerous and varying factors, which mainly originate from our social and cultural background. Nevertheless, a common and well- established factor that influences the shaping of music appraisal judgments is the level of familiarity with a musical excerpt. It has been demonstrated that even mere exposure to a particular stimulus can manipulate the degree of liking. In general, familiar stimuli tend to be more liked by subjects, although; over familiarization can result into negative judgments. This evidence implies an inverted U-shaped function of liking with respect to exposure. Unfamiliar stimuli are mapped to a high arousal potential and as familiarization takes place this potential decreases, causing liking to increase until an optimal level. Aspects of the latter effects in response to musical stimuli, and especially the U-shaped model, have been corroborated by several studies (Schellenberg et al.,2008, Hunter Schellenberg.,2010). Probing further, liking for music is mainly interpreted in terms of affective experiences. As music is a dynamic process, the content of which fluctuates

over time, the affective phenomena that it elicits, and, consequently, appraisal judgments, should also be considered as such processes (Bachorik et al.,2009). Since familiarity influences the shaping of music preference, it is expected to affect the unfolding of appraisal events over time, i.e., such events may peak in different time intervals during listening, depending on the familiarity with each excerpt. Research on the time course of affective responses to music has gained much interest over the past years. Musical excerpt is long can be categorized by subjects, with respect to their emotional content, in a similar way as much longer excerpts. Based on two- dimensional emotional ratings on average, a listener requires 8.31s of music before initiating an emotional judgment. The classification of electroencephalogram (EEG) responses to liked/disliked music was investigated by the broad researchers. Time-frequency (TF) analyses, energy-related features were estimated from EEG signals acquired from subjects during the listening of 15s long musical excerpts (Bachorik et al.,2009). Best classification accuracy (CA) {86.52($\pm$0.76)%} obtained by the combination of features from the beta (13-30 Hz) and gamma EEG bands (31-49 Hz), estimated in the total duration of music listening (0-15 s). Similar, but slightly lower, CA {84.94($\pm$0.94)%} achieved using the same features estimated in the time interval of 8-15 s after each excerpt onset. The EEG data set acquired can be reanalyzed by implementing a time-windowing approach for feature extraction, which is based on TF analysis (Hadjidimitriou and Hadjileontiadis.,2012).

3. PROPOSED WORK

3.1 Times-Frequency Distributions (TFDs)

A TFD constitutes a two-dimensional spectral representation of a signal in the time and frequency domains. There are two major classes of TFDs, i.e., the atomic decompositions (also referred as linear TF representations) and the quadratic TF representations. The first class involves a linear decomposition of the signal in elementary components, namely atoms, in order for the TFD to be acquired. An eminent member of this class is the short-time Fourier transform (STFT), which involves the prewindowing of a signal $x(T)$ around a time instant t and the subsequent calculation of the

Fourier transform for each t. The squared modulus of the STFT defines the spectrogram (SPG), which represents the energy distribution of the signal in the TF plane:

$$SPGx(t, f) = \left| \int_{-\infty}^{+\infty} x(\tau) h^*(\tau - t) e^{-j2\pi fr} d\tau \right|^2$$

where h*(T-t) represents the short-time analysis window (* denotes the complex conjugate). The major drawback of the STFT, and consequently of the SPG, is the trade-off between time resolution and frequency resolution. In particular, the time-window should be relatively small for a satisfactory time resolution, while, on the other hand, a good frequency resolution is acquired through a narrow- band filter, i.e., a long time window.

The second class of quadratic TF representations includes direct distributions of the energy of a signal in the TF plane (Auger et al.,1996, Boashash.,2003). Members of the latter class, which verify the properties of time and frequency covariance, form a subset of representations that is referred to as the Cohen's class (Cohen.,1989) and possess the following general expression:

$$C(t, f) = \int_{-\infty}^{+\infty} e^{j2\pi\xi(s-t)} g(\xi, \tau) x(s + \tau/2) x^*(s - \tau/2). e^{-j2\pi f\tau} d\xi ds d\tau$$

where t, f denote the time and frequency, respectively, x(t) is the signal, and $g(\xi, T)$ is the kernel function. The kernel is a function of the delay T and the doppler ξ in the so-called ambiguity plane. The major drawback of the quadratic TFDs is the appearance of cross terms (interferences), which arise from the quadratic superposition principle that the latter distributions involve. Certain members, namely reduced interference distributions (RIDs), of Cohen's class target at reducing these terms by a careful selection of the kernel function (Zhao et al.,1990). An eminent member of the RIDs group is the Zhao-Atlas-Marks (ZAM) distribution, which significantly reduces cross-terms by adopting a cone- shaped kernel function:

$$g(\xi, \tau) = h(\tau) \mid \tau \mid \frac{\sin(\pi\xi\tau)}{\pi\xi\tau}$$

where h(T) is a window function that leads to smoothing along the frequency axis. Thus, the following expression which defines the ZAM distribution can be obtained:

$$\int_{-\infty}^{+\infty} h(\tau) \left[\int_{t-|r|/2}^{t+|r|/2} x(s + t/2) x^*(s - t/2) ds \right] e^{-j2\pi fr} d\tau$$

Finally, a third method for the TF representation of a signal, which does not belong to the aforementioned classes, is the Hilbert-Huang spectrum (HHS) (Huang et al.,1998). The latter method consists of the empirical mode decomposition (EMD) of a signal into intrinsic mode functions (IMFs) and the subsequent application of the Hilbert transform on each IMF, in order for an analytic signal to be obtained:

$$Z_k(t) = IMF_k(t) + jIMF_k^H(t) = A_k(t) e^{j\theta_k(t)}$$

where $IMF_k(t)$ denotes the kth IMF, $IMF_k^H(t)$ is its Hilbert transform, $A_k(t)$ represents the amplitude of $Z_k(t)$ represents the instantaneous phase. Due to their generating process, IMFs constitute narrow band components and, therefore, a meaningful instantaneous frequency can be estimated by the derivative of $\theta_k(t)$ i.e., $f_k(t) = \frac{1}{2\pi}\frac{d\theta_k}{dt}$. The three-dimensional plot of the squared amplitude $A_k^2(t)$ for all IMFs, with respect to time and instantaneous frequency $f_k(t)$, forms the HHS-based TFD of a signal.

3.2 Time Frequency based Feature Extraction

Features were extracted from TFDs of EEG signals. Feature estimation was based on the concept of event-related desynchronization (ERD) and synchronization (ERS). The ERD/ERS theory postulates that event-related activity could be seen as a proportional change of the EEG spectral energy in relation to a reference period, placed some seconds before the stimulus onset (Pfurtscheller and Lopes da Silva.,1999). Additionally, such phenomena should be investigated in discrete EEG frequency bands as they arise from the intrinsic properties of neuronal populations functioning as resonating systems in specific frequency ranges (Niedermeyer and Lopes da Silva.,1998). Thus, after the estimation of the TFD TF [t,f] of an EEG epoch corresponding to the stimulation period, from recording channel i and experimental trial j, feature F is computed in a time window w_n and in the frequency band (f_b) of interest as

$$F^{fb,wn} = \frac{A^{fb,wn} - R^{fb}}{R^{fb}}$$

The quantity A^{f_b,w_n} is computed as the average TFD amplitude over both frequency and time, i.e.,

$$A^{fb,wn} = \frac{1}{N_{wn}} \sum_t \frac{1}{N_{fb}} \sum_f TF[t,\ f],$$

where N_{w_n} and N_{f_b} denote the number of samples in each time window w_n and the number of samples in each frequency band f_b, respectively. In a similar manner, the quantity R^{f_b} is computed as

$$R^{fb} = \frac{1}{N_R} \sum_t \frac{1}{N_{fb}} \sum_f TF_R[t,\ f],$$

where $TF_R[t,f]$ is the TFD estimated from the EEG epoch recorded during a reference time interval that precedes the stimulus onset (corresponding to the same trial j and recording channel *i*), N_R is the number of samples in the latter interval, and N_{fb} is the number of samples in the same frequency band f_b. It must be noted here that the TFDs are computed in their discrete forms.

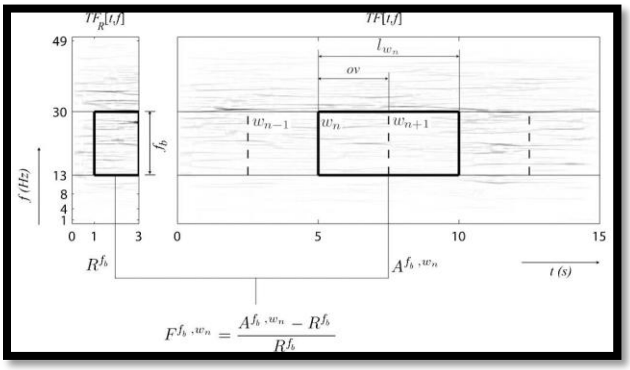

Fig. 1 Schematic representation of the time-frequency-based feature extraction procedure

As it is evident from the aforementioned, a time-windowing approach for feature extraction is adopted, i.e., feature F is computed from the TFD in different time windows w_n, covering the total duration of the stimulation period. Time windows are of length l_{w_n} and they are selected to be consecutive along the time axis with a percentage of overlapping (*ov*). Fig. 1 illustrates the feature extraction process described above (Hadjidimitriou and Hadjileontiadis.,2013). Subsequently, for each trial j, time window w_n and frequency band f_b, the feature vector (FV) is constructed as

$$FV_j^{fb,wn} = \left\{ F_{j,1}^{fb,wn}, \ldots, F_{j,i}^{fb,wn}, F_{j,Nc}^{fb,wn} \right\}$$

where *i* denotes the *i*th recording channel and N_c is the total number of channels.

3.3 Feature Estimation

The filtered EEG signals were subjected to the TF-based feature extraction procedure. For each trial, the stimulation period was the interval of music listening (15 s), while the reference period was the last 2 s of the resting interval preceding the beginning of the musical excerpt. The first second was omitted, as it succeeded subjects' rating, to avoid any resting brain activity. Feature estimation involved a time-windowing approach, i.e., features were estimated from the TFD in consecutive overlapping time-windows. Thus, the minimum value of the window length and the overlapping percentage between consecutive windows had to be set.

According to the literature, the time interval which subjects require to demonstrate an effective response or to assess the emotional content of a musical piece is in the order of seconds. As far as the EEG frequency bands are concerned, features were estimated from the beta (13-30 Hz) and gamma (31-49 Hz) EEG bands. The cut-off frequency for the gamma band was set to 49 Hz by taking into consideration the final sampling frequency of the recording device, i.e., 128 Hz, which allows for a bandwidth of 64 Hz, as well as the hardware and software-based online filtering which could distort EEG data in higher frequencies. The focus was placed upon the latter frequency bands as the combination of features derived from them led to statistically significant higher classification accuracies as compared to other EEG bands and, consequently, to the best discrimination between music appraisal responses. The TF-based feature extraction procedure was implemented as follows (Hadjidimitriou and Hadjileontiadis.,2013).

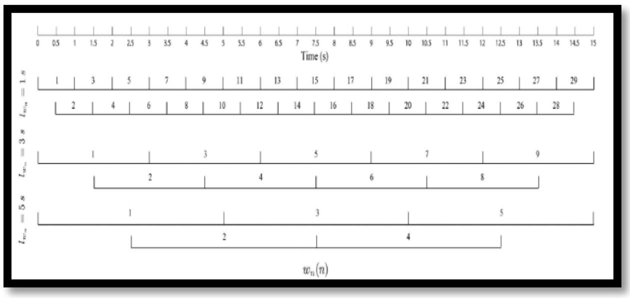

Fig. 2: Indexing (n) of the consecutive time windows wn, for each window length lwn = 1, 3, 5 s, during the course of music listening (15 s).

For each trial, the TFDs of the EEG epochs acquired during the stimulation and resting periods were computed for the 14 recording channels. Subsequently, the FV for each band was constructed. The final FV for the *jth* trial and the nth time window (*wn*) was formed by concatenating the FV s corresponding to the beta and gamma bands, i.e.,

$$FV_j^{wn} = \left\{ FV_j^{beta,wn}, FV_j^{gamma,wn} \right\}$$

For each channel and each of the two frequency bands, R^{f_b} was computed from the TFD corresponding to the resting interval using FV.

Feature F^{f_b,w_n},was computed using R^{f_b} and f_b, w_n , estimated from the TFD corresponding to the interval of music listening, in time window w_n and in each frequency band of interest. The complete FV set corresponding to time window wn consisted of the FV s computed for all trials of music listening. The total number of FV sets computed during the time course of music listening was $\frac{15-0.5l_w}{0.5l_w}$, equal the number of time windows; 15 and 0.5 denote the duration of the stimulation period in seconds and the fraction of overlapping between time windows, respectively. Fig. 2 depicts the number and indices (n) of time windows wn for each window length l_{wn} For example, 29 FV sets were computed during the time course of music listening using 1-s-long consecutive time windows with 50 percent overlap. For the computation of the TFDs, the three methods were employed, i.e., the SPG, the ZAM method, and the HHS. A 64-point fast Fourier transform (FFT) with a nonoverlapped Hamming window of 500 ms was used for the computation of the SPG. The nonoverlapped windows were selected to acquire completely disjoint energy estimations from the different EEG segments.

The length of the short-time window was set to 500 ms for the subsequent time-window segmentation, described above, to be implemented. The ZAM distribution was computed under a N x N TF resolution; N denotes the number of samples of the signal.

Smoothing was performed using Hamming windows of N=10- samples and N/4-samples for time and frequency (Auger et al.,1996, Boashash.,2003). Finally, the HHS was computed by implementing the EMD algorithm (Huang et al.,1998). Fig. 3 illustrates the produced TFDs from an EEG signal using the three TF representation methods (Hadjidimitriou and Hadjileontiadis.,2013). The aforementioned analysis and the subsequent classification procedure were conducted using Matlab R2011b (Mathworks Inc., Natick, MA).

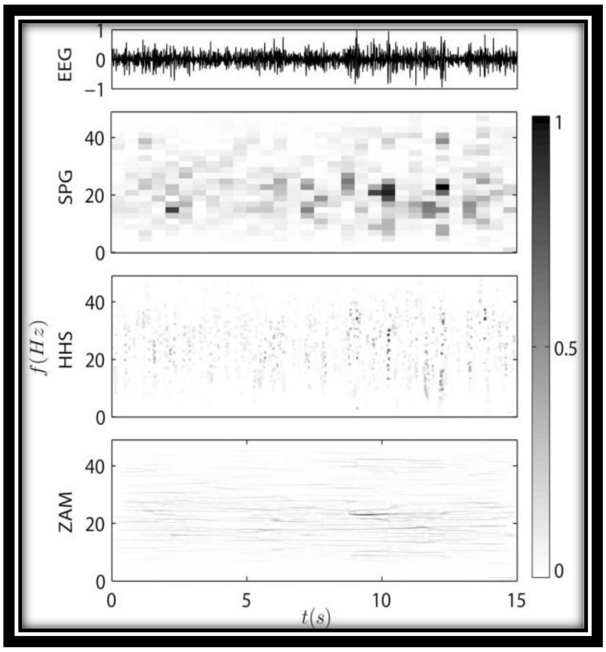

Fig. 3: Normalized time-frequency distributions (TFDs) of a 15-s-long EEG signal (first row) computed using the SPG (second row), the ZAM (third row), and the HHS (fourth row) methods.

3.4 Classification

The classification approach adopted in this study was user independent, i.e., classification was performed on the complete data set, created from all subjects' EEG responses. Initially, the estimated *FV*s were grouped in two classes, namely "like" and "dislike." Class "like" comprised of FV s corresponding to the trials in which subjects rated their liking for the musical excerpts as 5 (like very much) or 4 (like).

Consequently, class "dislike" comprised of FV s corresponding to the trials in which subjects rated their liking for the musical excerpts as 2 (do not like) or 1 (do not like at all). The FV sets were further categorized according to the level of familiarity. *FV*s corresponding to trials in which subjects rated their level of familiarity with the musical excerpt as 5 (very familiar) or 4 (familiar) were labelled as "familiar," while FV s corresponding to trials in which subjects rated their level of familiarity with the musical excerpt as 2 (unfamiliar) or 1 (totally unfamiliar) were labelled as "unfamiliar." As presented in Table 1, only few significant high correlations between subjects' liking and familiarity ratings for the 60 musical excerpts were observed, indicating that positive/negative appraisal responses were deduced from different excerpts per listener.

In short, three types of data sets were considered: 1) the FV sets corresponding to liking and disliking ratings of familiar excerpts (LDF), 2) the FV sets corresponding to liking and disliking ratings of unfamiliar excerpts (LDUF), and 3) the FV sets corresponding to liking and disliking ratings, regardless of familiarity (LD), i.e., the joint LDF and LDUF FV sets. The purpose of the binary classification procedure was to discriminate between liking and disliking responses under the parameter of familiarity.

4. ABOUT THE DATASET

The EEG data set, by (Hadjidimitriou and Hadjileontiadis.,2012), was acquired for the sake of clarity; a summarized description of the experimental procedure is given here. Nine subjects participated in an experiment during which they listened to 60 15-s-long nonvocal musical excerpts and 15 15-s-long excerpts of broadband noise, while their brain activity was recorded. Musical excerpts belonging to the four most common genres, i.e., rock-pop, electronic, jazz, and classical (15 excerpts per genre), were carefully selected to represent the intrinsic characteristics of each genre, e.g., colourful instrumentation for symphonic pieces, jazz scales, and characteristic instrument textures that tie with the definition of jazz. Consequently, a kind of uniform stimulation of excitement per genre was anticipated for each user, based on the objective characteristics of the music that essentially define the genre itself. The total number of trials was equal to the number of excerpts, i.e., 75 per subject and 9 x 75 = 675 in total. EEG signals were acquired using the Emotiv EPOC wireless recording headset (Emotiv Systems, Inc., San Francisco, CA) from 14 channels (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8), referenced to the common mode sense (CMS—left mastoid)/driven right leg (DRL—right mastoid) ground. The EEG data are acquired with an internal sampling frequency of 2,048 Hz by the recording device, and, subsequently, band-pass filtered in the range of 0.16- 85 Hz using hardware filters. All EEG signals were band-pass filtered offline in the frequency range of 1-49 Hz (Hadjidimitriou and Hadjileontiadis.,2012).

5. RESULTS AND DISCUSSIONS

The following are the performance metrics used to evaluate the performance of the two proposed classification methods namely liking & similarity rating method. The classification is performed by two types namely calm and excited. Sampling – 1 consists of 100 EEG signal input. The classification accuracy, sensitivity and specificity can be calculated using the following metrics.

- ✓ True Positive is the Calm Positive
- ✓ True Negative is the Calm Negative
- ✓ False Positive is the Calm Excited
- ✓ False Negative is the Calm Positive Negative
- ✓ Neutral denotes Negative Positive [Not considered]

Table 1. depicts the true positive, true negative, false positive, false negative, sensitivity, specificity and accuracy performance for sampling 1 of the algorithms such as Linear Discriminant Analysis (LDA), Linear Support Vector Machine (LSVM), Radial Basis Function based Support Vector Machine (RBF – SVM), the proposed work named Liking and Familiarity Rating (L & F). It can be clearly understood that the proposed work L & F provides better results.

Table 1. Comparative Analysis of TP, TN, FP, FN, Sensitivity, Specificity, Accuracy

Method	TP	TN	FP	FN	Sensi (%)	Speci (%)	Accu (%)
LDA	48	15	10	27	35.7	64.0	63.4
LSVM	65	13	12	10	56.5	86.7	78.8
RBF-SVM	61	13	12	14	48.1	81.3	74.7
L & F	85	10	4	1	90.9	98.8	95.4

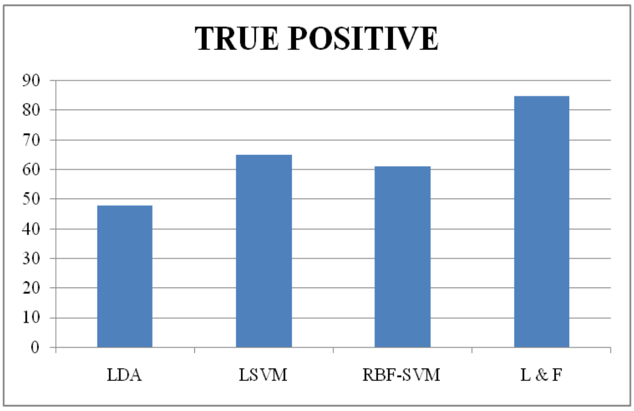

Fig 4. True Positive Analysis

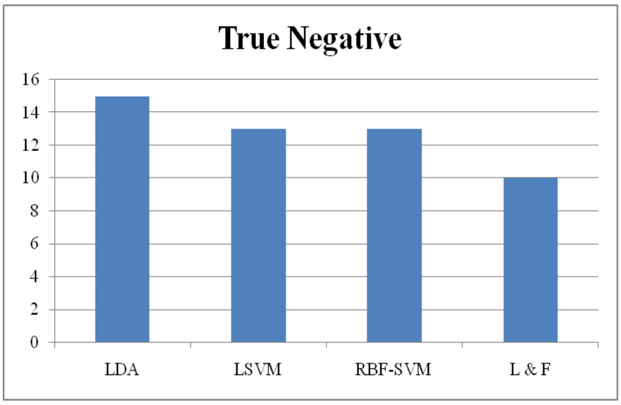

Fig 5. True Negative Analysis

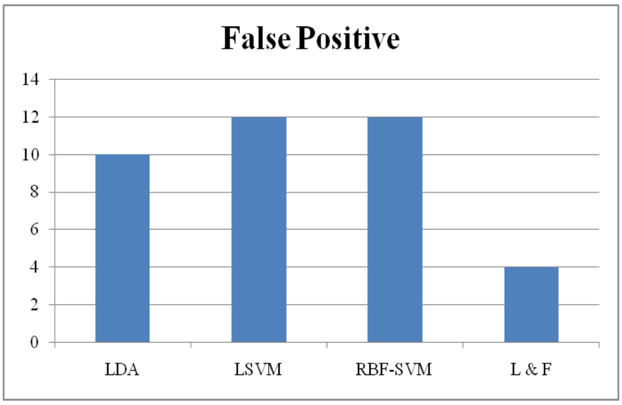

Fig 6. False Positive Analysis

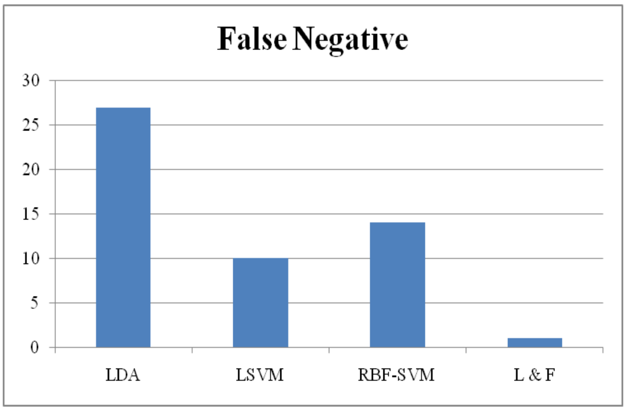

Fig 7. False Negative Analysis

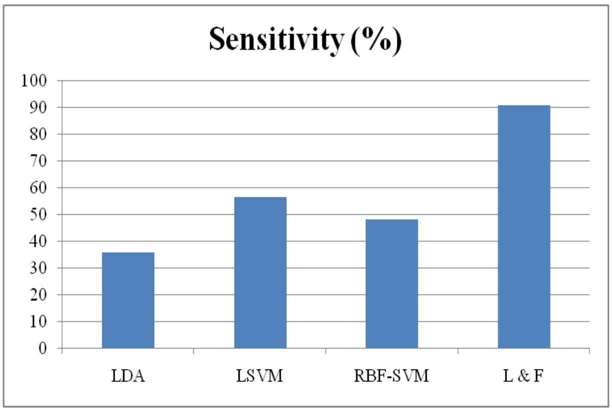

Fig 8. Sensitivity Analysis

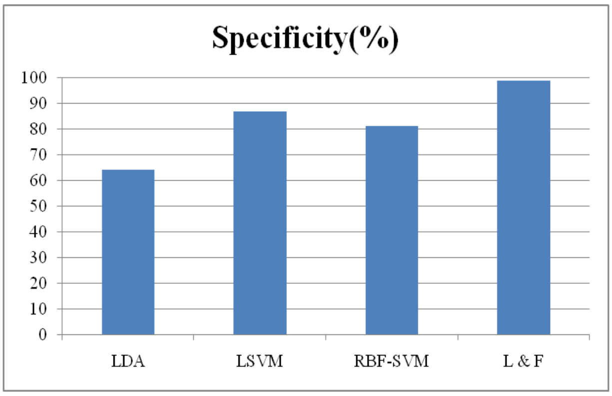

Fig 9. Specificity Analysis

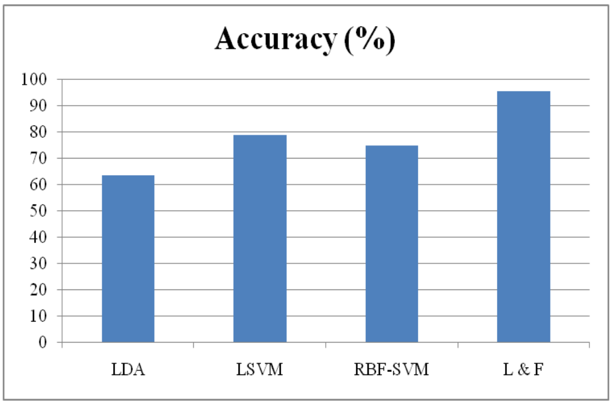

Fig 10. Accuracy Analysis

From the above figures it is very clear that L &F performs in an outstanding manner, which gives better sensitivity, specificity and accuracy

6. CONCLUSIONS

The most famous strategy used for perusing mind movement is electroencephalography (EEG). Electroencephalography is the neurophysiologic estimation of the electrical action of the cerebrum by recording from anodes put on the scalp, or in the exceptional cases on the cortex. The proposed linking and familiarity rating method classified the music, video assessment responses of EEG-Signal. The performance metrics namely true positive, true negative, false positive, false negative, sensitivity, specificity and classification accuracy are chosen for the proposed classifier. The simulation results showed that the proposed classifier achieves 95.4 % accuracy which is better than other methods.

7. REFERENCES

[1] Adams, Bahr, Moreno, "Brain Computer Interfaces: Psychology and Pragmatic Perspectives for the Future", AISB 2008 Convention, Aberdeen, Scotland.

[2] Auger, Flandrin, Goncalves, Lemoine, Time-Frequency Toolbox Tutorial. Centre Nat'l de la Recherche Scientifique (CNRS)/Rice Univ., 1996.

[3] Bachorik, Bangert, Loui, "Emotion in Motion: Investigating the Time-Course of Emotional Judgments of Musical Stimuli," Music Perception, vol. 26, no. 4, pp. 355-364, 2009.

[4] Boashash, "Theory of Quadratic TFDs," Time Frequency Signal Analysis and Processing: A Comprehensive Reference, B. Boashash, ed., pp. 59-82, Elsevier, 2003.

[5] Borstein, "Exposure and Affect: Overview and Meta-Analysis of Research 1968-1987," Psychological Bull., vol. 106, no. 2, pp. 265-289, 1989.

[6] Cohen, "Time-Frequency Distributions—A Review," Proc. IEEE, vol. 77, no. 7, pp. 941-981, July 1989.

[7] Hadjidimitriou and Hadjileontiadis, "Towards an EEGBased Recognition of Music Liking Using Time-Frequency Analysis," IEEE Trans. Biomedical Eng., vol. 59, no. 12, pp. 3498-3510, Dec. 2012.

[8] Hadjidimitriou, Hadjileontiadis,"EEG-Based Classification of Music Appraisal Responses Using Time-Frequency Analysis and Familiarity Ratings", IEEE Transactions On Affective Computing, Vol. 4, No. 2, April-June 2013

[9] Huang, Shen, Long, Wu, Shih, Zheng, Yen, Tung, Liu, "The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis," Proc. Royal Soc. London. Series A: Math., Physical and Eng. Sciences, vol. 454, no. 1971, pp. 903-995, 1998.

[10] Hunter Schellenberg, "Interactive Effects of Personality and Frequency of Exposure on Liking for Music," Personality Individual Differences, vol. 50, no. 2, pp. 175-179, 2010.

[11] Niedermeyer, Lopes da Silva, "Electroencephalography: Basic Principles, Clinical Applications and Related Fields", Williams and Wilkins, 1998.

[12] Pfurtscheller, Lopes da Silva, "Event-Related EEG/MEG Synchronization and Desynchronization: Basic Principles,"Clinical Neurophysiology, vol. 110, pp. 1842-1857, 1999.