Emotion driven mood enhancing multimedia recommendation system using physiological signal

Sindhu N¹, Jerritta S², Anjali R¹
¹Government Engineering College, Thrissur
²VISTAS, Chennai
E-mail: sindhun@gecr.ac.in, jerritta.se@velsuniv.ac.in, anjalirajan.7994@gmail.com

Abstract. This paper exhibits a proposal for developing a multimedia recommendation system that is driven by the emotion of the user. The multimedia system is designed in such a way that the user will not have to manually browse for the audio/video to be played, it will be played automatically based on the emotion. The database comprises ECG signals collected from DECAF. Anger and sadness, emotions contributing to negative mood are considered. Discrete Wavelet Transform (DWT) was used for feature extraction and classification accuracy was determined using kNN, SVM. It was found that kNN classifies better than SVM. Real-time signals of one single healthy subject were collected based on a protocol built. Multimedia recommendation system operates automatically based on the classified emotion.

Keywords: Emotion, ECG, DECAF, classification, multimedia recommendation system.

1. Introduction

Emotions play a major role in deciding the demeanor of a person, how he reacts to situations around him, to the various tasks assigned to him etc. Effective understanding of a person’s emotion can therefore contribute greatly in making his life easy. Affective computing or human computer interaction using human emotions is one of the emerging fields now. Emotions of a person can be analysed from their facial expression, physiological changes like blood pressure variation, change in heart rate, respiration rate, voice modulation, etc. Every psychological reaction has an underlying physiological basis to it. Analysis of physiological signals like electrocardiograph (ECG), electroencephalograph (EEG), photoplethysmograph (PPG), galvanic skin response (GSR), skin temperature (SKT), blood volume pulse (BVP), respiration, electromyography (EMG) etc are hence extensively used for recognizing emotions. Physiological signals being controlled by autonomic nervous system offers the advantage that these cannot be masked [2]. Emotion detection from ECG signals is gaining attention, and the process of emotion detection from ECG signals has the following steps: 1) data acquisition, 2) data pre-processing, 3) feature extraction and 4) classification [3]. Entwining emotion detected from ECG and a multimedia recommendation system operating based on the emotion would help develop a multimedia recommendation system that learns the emotion of a user from the signals obtained via wearable physiological sensors [1].
2. Related works

Alexandros Zenonos et al. [3] proposes a mood recognition framework using physiological signals recorded using wearable sensor devices and devices measuring motion signals. Physiological signals such as ECG, PPG, Pulse Wave Transit Time (PWTT), skin temperature and 3 axial acceleration of each subject was recorded using Toshiba Silmee wearable smart wrist band. Simultaneously, the subjects were also asked to submit their moods using the HealthyOffice app installed on their smart phones.

In [1], Andreas Haag et al. proposes use of bio sensors, ProComp+ in emotion identification. Simultaneously the physiological signals were collected using ProComp+ bio sensor which is a collection of different bio sensors. The physiological signals like ECG, BVP, skin temperature, respiration, skin conductivity and EMG were collected from different parts of the body. In [2], Deger Avata et al. proposes emotion based music recommendation system that uses physiological signals GSR, PPG and EEG for music recommendation. Along with time domain features like mean, median, first degree moments, standard deviation and root mean square standard deviation, non-linear features like skewness and kurtosis were also obtained. In [3] Schleusing et al developed a sensing garment and portable sensors that can measure multiple physiological signals like ECG, speech and respiration of patients affected by bipolar disorders. The aim of the research was early indication of bipolar disorder and it was concluded that physical changes and mood variations have a high correlation.

Arya et al. [4] performed an analysis on the performance of different emotion elicitation methods was studied in 25 people including both male and female in the age group of 21-25 based on their ECG signals recorded using LabChart. DWT was used for extracting the heart rate variability (HRV) features and support vector machine (SVM) for classification of the emotion as happy or sad. Adheena et al. [5] performed anxiety detection. Anxiety contributing to negative emotions and other cardiovascular diseases (CVD) was detected for 15 chosen subjects. ECG signal of the subject was recorded and the features (QRS complex) were detected using the Pan Tomkins algorithm. The correlation between inter beat interval/R-R interval and anxiety, relation between anxiety arousal and baseline condition were detected. Classification was done using SVM and Kalman Filter. In [6], Priyanka et al. calculated the performance comparison of two time frequency analysis (Hilbert Huang Transform (HHT) and DWT) from database offered by DECAF. The wavelets chosen in the work ranges from Db6, Db7, sym8 to coif5. 10 subjects out of the 30 available in the database were tested and their happy and sad emotions were detected. Amani Albraikan et al. [7] tried to develop a user independent emotional model that would increase the accuracy of existing emotional learning models. Datasets were chosen using customised Empatica E4 sensor and MAHNOB for measuring heart rate using PPG, BVP, electro dermal activity (EDA), axial accelerometer and skin temperature using infrared temperature sensor. In [8] Paradiso et al. developed a mental health care monitoring system (PSYCHE) for helping people affected with bipolar disorders. A sensing garment was developed that was capable of recording physiological signals like ECG, GSR, EMG, EEG, Electrooculograph (EOG), blood pressure, body temperature, respiration. The features extracted were heart rate (HR), HRV, breathing rate and breathing amplitude. The overall aim of the project was development of a personalized sensing garment [9].

3. Proposed method

The proposed method uses ECG signal for emotion identification. After classification of the emotion, the multimedia recommendation system operates automatically. The recommendation of multimedia will be based on the past preferences of the user, which has already been collected
and saved as database. The block diagram of proposed work is shown in figure 1 and figure 2 below.

![Figure 1. Block diagram of proposed system](image1)

![Figure 2. Block diagram of proposed multimedia recommendation system](image2)

3.1. ECG signal acquisition

The database used is ECG signals of negative mood emotions angry and sad from DECAF [10], a multimodal dataset for decoding user physiological responses to affective multi-media content. This database consists of emotional responses of 30 participants, in which the emotion is evoked by 40 one-minute music video segments and 36 movie clips. ECG signals were recorded using three sensors attached to the participant. Two electrodes were placed on the wrist, and a reference was placed on a boney part of the arm (ulna bone).

3.2. Pre-processing

The raw ECG signals collected were further loaded pre-processed to remove artefacts like power-line interference removed using Notch filter designed at 50 Hz, followed by low pass Butterworth filter designed at and high pass Butterworth filters so as to obtain the relevant frequencies in the range of 0.04-0.15 Hz.

3.3. Feature extraction

R peak detection of ECG signals was done using Pan Tompkins Algorithm. HR and HRV was determined and the 14 level decomposition of HRV using DWT was performed so as to obtain the required coefficients for emotion detection. Mean, median, standard deviation, maximum value, minimum value, power and ratio of each signal was calculated for relevant coefficients 11, 12, 13 and 14.

3.3.1. Pan Tompkins Algorithm:
Pan Tompkins algorithm is used for detection of QRS complex from real time ECG signals. The signal input first goes into a cascaded high-pass and low-pass filter, performing action of band pass filter for removal of noise. The next stage is the stage of differentiation, where derivative is taken so as to identify the peaks. After taking derivative of
the signal, then squaring is done so as to intensify the slope of the frequency response curve of the derivative and helps restrict false positives by T waves. Further the moving window integrator provides information about the slope and width of QRS complex. The algorithm has three phases including learning phase 1, learning phase 2, and finally detection phase. Learning phase 1 includes the initialization of detection thresholds based upon signal and noise peaks detected during the process. Learning phase 2 includes initialization of RR-interval average and RR-interval limit values using two heartbeats. Finally, the detection phase is responsible for the recognition and produces a pulse for each QRS complex. The thresholds and other parameters of the algorithm are adjusted periodically to adapt to changing characteristics of the signal.

3.3.2. **DWT:** It is a time frequency analysis method which has high frequency resolution. Since emotional data is present in low frequency and high frequency of signal, to extract these components a 14-level decomposition is done using wavelet transform. 11th, 12th, 13th and 14th level coefficients are used for feature extraction. Low frequency (LF) power, high frequency (HF) power, total power and ratio of HF power to LF power are the features extracted.

3.4. **Classification**  
3.4.1. **kNN:** The output of kNN classification is a class membership. Initially, k value, a positive integer, typically a small value is chosen. The object will be assigned to the class into the most common in which its k nearest neighbors belong.

3.4.2. **SVM:** The goal of support-vector machine is to construct a hyperplane or set of hyperplanes in a higher dimensional space. It can be used for classification, regression, or other tasks such as outliers detection. They are supervised learning models with associated learning algorithms for analyzing data used for classification and regression analysis. All the training examples will be marked as belonging to one or the other categories. SVM model represents the examples as mapped points in space. These points of the separate categories are divided by a clear gap that is as wide as possible. Then based on the side of gap they fall, new examples are predicted to belong to a category.

3.5. **Protocol Development**  
Signal acquisition can be done only on the basis of an accurate protocol developed. The protocol was developed by asking the subject regarding the videos/audios and images that would make them happy, sad and angry. The protocol was designed for a duration of 4 minutes 30 seconds and displayed to the subject using mobile phone and headset. Signal acquisition was done in 2 sessions in a day.

3.6. **Multimedia recommendation system**  
The designed framework operates in such a manner that is based on the emotion, a video or audio starts playing. The video or audio is collected based on the interests and preferences of the subject whose signal is collected and analysed. These are those video/audio clips that are intended to make the subject happy. The traditional recommendation frameworks and systems operate on the basis of users search history, most played, browsing history etc.

4. **Results**  
The whole process is executed in Windows 10 PC with 64-bit processor and 4 GB internal RAM. Program is executed using MATLAB R2016a. The datas are collected from DECAF [12] database, a multimodal data set for decoding user physiological responses to affective multimedia.
content. Signals of 25 subjects were taken for database creation and 5 subjects were tested. Pre-processed output of raw data signals is shown in figure 3. After feature extraction the features were classified using SVM and kNN classifier. Figure 4 shows classification of emotion after feature extraction. It was found that the classification accuracy of kNN is higher compared to SVM. Real time signal acquisition was done using Shimmer3 Consensys ECG Development kit for Multimedia Recommendation System. Shimmer device used for data acquisition is shown in Figure 5 and figure 6 shows electrode placement PC connection with Shimmer3 Consensys Development Kit. Figure 7 shows each stage of preprocessing of real time signal and the final output of multimedia recommendation system is shown in figure 8.

Figure 3. Pre-processing of raw data signals from DECAF

Figure 4. Classification of emotion after feature extraction
Figure 5. Wearable Shimmer 3 device and electrode placement

Figure 6. PC connected with Shimmer3 Consensys Development Kit

Figure 7. Pre-processing of real time signal acquired
Performance analysis of both the database and real time signal acquired was calculated using the following equations.

\[
\text{Accuracy} = \frac{\text{TruePositive} + \text{TrueNegative}}{\text{TruePositive} + \text{TrueNegative} + \text{FalsePositive} + \text{FalseNegative}} \quad (1)
\]

\[
\text{Sensitivity} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}} \quad (2)
\]

\[
\text{Specificity} = \frac{\text{TrueNegative}}{\text{TrueNegative} + \text{FalsePositive}} \quad (3)
\]

\[
\text{Precision} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalsePositive}} \quad (4)
\]
Table 1. Performance analysis values from database

| Emotion | Accuracy | Specificity | Sensitivity | Precision |
|---------|----------|-------------|-------------|-----------|
| Angry   | 0.70     | 0.5625      | 0.75        | 0.7058    |
| Sad     | 0.70     | 0.8571      | 0.6428      | 0.6923    |

Table 2. Performance analysis values from real time signal

| Emotion | Accuracy | Specificity | Sensitivity | Precision |
|---------|----------|-------------|-------------|-----------|
| Angry   | 0.75     | 0.3333      | 0.6667      | 1.0       |
| Sad     | 0.75     | 0.2         | 1.0         | 0.50      |

The result was found as shown in tables, Table 1 and 2. The performance value shows that specificity is 88 percent for true negative which implies that emotion sad is correctly identified as negative state with high accuracy.

5. Conclusion

In current trend of smart appliances around us, emotion also plays a vital role in making human computer interaction more effective and easy. Affect Computing will be taken to the next level when user emotion is also incorporated into its operation. The work done in this regard would certainly contribute positively in using physiological signals, specifically ECG for detection of affect state. Development of a fully automatic system will indeed be one of the greatest contributions that can be made to the domain of smart appliances that humans use in daily life. The work can be further extended by collecting more number of real time signals, performing online validation of the system, using more number of classifiers, use of other physiological signals as well for emotion detection and by developing a fully automatic system for the purpose.

6. References

[1] Deger Ayata, Yusuf Yaslan, and Mustafa E. Kamasak, Emotion Based Music
[2] Carlos Lara-A lvarez, Hugo Mitre-Hernandez, Juan J. Flores, and Humberto Pe rez-Espinosa, Induction of Emotional States In Educational Video Games Through A Fuzzy Control System, Ieee Transactions On Affective Computing, Vol. 0, No. 0, August 2018.
[3] Alexandros Zenonos, Aftab Khan, Georgios Kalogridis, Stefanos Vat- sikas, Tim Lewis, Mahesh Sooriyabandara, HealthyOffice: Mood Recogni- tion At Work Using Smartphones and Wearable Sensors, The Second IEEE International Workshop on Sensing Systems and Applications Using Wrist Worn Smart Devices, 2016.
[4] O. Schleusing, Ph. Renevey, M. Bertschi, St. Dasen, J.-M. Koller, R. Paradiso, Monitoring physiological and behavioral signals to detect mood changes of bipolar patients, 978-1-4244-9445-3/11/26.00 2011 IEEE.
[5] Paul Baskett, Yi Shang, Michael V. Patterson, Timothy Trull, Towards A System for Body-Area Sensing and Detection of Alcohol Craving and Mood Dysregulation, The 10th Annual IEEE Consumer Communications and Networking Conference Demos.
[6] R. Paradiso, A. M. Bianchi, K. Lau and E. P. Scilingo PSYCHE: Personalised Monitoring Systems for Care in Mental Health, 32nd Annual International Conference of the IEEE EMBS Buenos Aires, Argentina, August 31 - September 4, 2010.
[7] Rafael A. Calvo, Senior Member, IEEE, and Sidney DMello, Member, IEEE Computer Society, Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications, Ieee Transactions On Affective Computing, Vol. 1, No. 1, January-June 2010.

[8] Dimitris Giakoumis, Dimitrios Tzovaras, Member, IEEE, Konstantinos Moustakas, Member, IEEE, and George Hassapis, Senior Member, IEEE, Automatic Recognition of Boredom in Video Games Using Novel Biosignal Moment-Based Features, Ieee Transactions On Affective Computing, Vol. 2, No. 3, July-September 2011.

[9] Andreas Haag, Silke Goronzy, Peter Schaich, Jason Williams, Emotion Recognition Using Bio-Sensors: First Steps Towards an Automatic System, E. Andr et al. (Eds.): ADS 2004, LNAI 3068, pp. 3648, 2004. Springer-Verlag Berlin Heidelberg 2004.

[10] Priyanka Chettupuzhakkaran, Sindhu N, Emotion Recognition from Physiological Signals using Time-Frequency Analysis Methods, Proceedings of 2018 International Conference on Emerging Trends and Innovations in Engineering and Technological Research (ICETIETR).

[11] Rangaraj M Rangayyan, Biomedical Signal Analysis - A Case Study Approach, WILEY INTERSCIENCE, John Wiley Sons, 2002.

[12] Mojtaba Khomami Abadi, Ramanathan Subramanian, Seyed Mostafa Kia, Paolo Avesanì, Ioannis Patras and Nicu Sebe, DECAF: MEG based Multimodal Database for Decoding Affective Physiological Responses, IEEE Transactions on Affective Computing, Vol. 6, No. 3, July-September 2015.