



# Emotion recognition in response to traditional and tactile enhanced multimedia using electroencephalography

Aasim Raheel<sup>1</sup> · Syed Muhammad Anwar<sup>1</sup> · Muhammad Majid<sup>1</sup> 

Received: 15 January 2018 / Revised: 4 October 2018 / Accepted: 19 November 2018 /  
Published online: 23 November 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

The goal of this study is to enhance the emotional experience of a viewer by using enriched multimedia content, which entices tactile sensation in addition to vision and auditory senses. A user-independent method of emotion recognition using electroencephalography (EEG) in response to tactile enhanced multimedia (TEM) is presented with an aim of enriching the human experience of viewing digital content. The selected traditional multimedia clips are converted into TEM clips by synchronizing them with an electric fan and a heater to add cold and hot air effect. This would give realistic feel to a viewer by engaging three human senses including vision, auditory, and tactile. The EEG data is recorded from 21 participants in response to traditional multimedia clips and their TEM versions. Self assessment manikin (SAM) scale is used to collect valence and arousal score in response to each clip to validate the evoked emotions. A t-test is applied on the valence and arousal values to measure any significant difference between multimedia and TEM clips. The resulting p-values show that traditional multimedia and TEM content are significantly different in terms of valence and arousal scores, which shows TEM clips have enhanced evoked emotions. For emotion recognition, twelve time domain features are extracted from the preprocessed EEG signal and a support vector machine is applied to classify four human emotions i.e., happy, angry, sad, and relaxed. An accuracy of 43.90% and 63.41% against traditional multimedia and TEM clips is achieved respectively, which shows that EEG based emotion recognition performs better by engaging tactile sense.

**Keywords** Emotion recognition · Multimedia · Tactile enhanced multimedia · Classification · Electroencephalography

---

✉ Muhammad Majid  
m.majid@uettaxila.edu.pk

Aasim Raheel  
asim.raheel@uettaxila.edu.pk

Syed Muhammad Anwar  
s.anwar@uettaxila.edu.pk

<sup>1</sup> Department of Computer Engineering, University of Engineering and Technology, Taxila, Pakistan

## 1 Introduction

Traditional multimedia content is a combination of video and audio components, which engage two human senses i.e., vision and auditory and is termed as bi-sensorial media [43]. It is used in every aspect of life but has found more significance in fields such as education, learning, and promoting innovation in business. Currently, multimedia content is limited in providing an immersed sense of reality and fails to deliver a real environmental sensation to the observers [44]. Humans perceive the real world environment using five different senses i.e., vision, auditory, gustatory, tactile, and olfactory. The traditional multimedia content fails to deliver certain sensations for example, a person walking through the street in a movie scene and experiencing the effects of air or humidity. Such content is termed as multiple sensorial media (mulsemdia) [15], which is generated by synchronizing specific devices with the audio-visual content to artificially generate aromas, humidity or air effects to engage additional senses of the viewer. Mulsemdia enhances the human experience of viewing digital content and provides a more realistic environment. Tactile enhanced multimedia (TEM) engages sense of touch or feel in addition to vision and auditory, enriching the human experience of viewing digital content. In a movie scene, olfaction effect is added by synchronizing devices such as a scent dispenser by identifying the timestamps [30, 31]. The quality of experience of users in terms of olfaction and tactile is evaluated by subjective self-reported tests. These are valuable in identifying the quality of experience, but sometimes affected by corroboration [7], where people might not respond with their true feelings. The quality of experience in response to multimedia content needs to be evaluated by using the brain response [2].

Emotions are mental states or internal feelings of the user and are evoked in response to external stimuli. Human emotions are used to communicate and interact with the outside world as a non-verbal language and can play an important role in rational decision making. Human emotions are generally categorized into six different types i.e., happy, sad, fear, disgust, joy, and anger. Speech signals [4], body gestures [11], and facial expressions [19] are conventional methods to recognize human emotions. Automatic human emotion recognition using facial expressions is a demanding and interesting area in the field of computer vision [13, 34]. These methods have challenges such as camera positioning and privacy. In computer vision based systems, a user has to be in the view of a camera in order to record his/her facial expressions. Experiments have used a fixed distance of either 60 cm or 80 cm from the stimuli and a head and chin rest to stabilize the head position [12]. Hue, saturation and intensity is also adjusted to detect facial expressions of an individual [26].

Recently, human emotion recognition based on physiological signals like skin conductance, electrocardiogram (ECG), and electroencephalography (EEG) in response to different stimuli has gained attention [20–22, 33]. Among these, EEG is widely used in emotion recognition because emotion based signals are generated from the limbic system that directs our attention [8]. Emotion is a complex cognitive function generated in the brain and associated with several different oscillations [6]. Emotions are event related potentials that directly measure cognitive states and are induced as fluctuating voltages during an event [35]. A symmetric electrode pair is used to calculate the alpha power asymmetry, which is commonly used for human emotion recognition [3, 17, 38]. Other spectral changes in EEG and regional brain activations are also related to human emotional states i.e., alpha and theta-power changes at the right parietal lobe [1, 36], beta-power asymmetry at the parietal region [39], and gamma spectral differences at the right parietal regions of the brain [5].

Bio-inspired multimedia research is primarily focusing on evaluating digital content and enriching human experience. The digital content is evaluated in response to different stimuli such as audio [8, 27], video [41], and images [32]. Different time and frequency domain EEG features have been used to recognize emotions in recent studies by using support vector machine (SVM) as a classifier [2]. In [27], two classifiers are compared i.e., SVM and multi-layer perceptron by using frequency domain features and music as a stimuli. In [8], three different classifiers are compared by extracting time and frequency domain features in response to music. In [41], emotions are recognized in response to videos by using frequency domain features and SVM as a classifier. To the best of our knowledge, no emotion recognition study based on EEG signal in response to TEM is available.

In this paper, TEM is generated by synchronizing an electric fan and a heater with the audio-visual content after identifying timestamps in a scene to add an effect of cold and hot air respectively. The generated TEM content is then evaluated in comparison to the traditional multimedia content by assessing cognitive states of the observer in terms of emotion. Self-assessment manikin (SAM) scale is used to label the evoked human emotion in response to both TEM and traditional multimedia content. The emotions are recognized in response to TEM and traditional multimedia using EEG signals. Twelve time domain features are extracted from the acquired EEG data to classify four different emotions i.e., happy, angry, sad, and relaxed using SVM classifier. The main contributions of this paper are,

1. TEM clips are generated by synchronizing a fan and a heater to add cold and hot air effects in traditional multimedia content.
2. Statistical significance of TEM in terms of valence and arousal is established.
3. Emotion recognition in response to traditional multimedia and TEM content using time domain features of acquired EEG signals are reported.

The rest of the paper is organized as follows. Section 2 presents a brief overview of literature related to emotion recognition using EEG signals. Generation and synchronization of TEM is discussed in Section 3. The proposed human emotion recognition in response to TEM and traditional multimedia content using EEG signals is presented in Section 4. Emotion recognition results using EEG in response to traditional and TEM are discussed in Section 5 followed by conclusion in Section 6.

## 2 Related works

In literature different emotion recognition techniques using EEG signals have been proposed. The major steps involved in EEG based human emotion recognition are data acquisition, preprocessing, feature extraction, and classification. For EEG based emotion recognition, data are acquired in response to different stimuli that engage human senses like auditory, vision, olfaction, and tactile. Different methods of EEG based emotion recognition in response to audio music that only involve auditory sense have been proposed [8, 27]. In [27], machine learning based approach to recognize four different emotions in response to music listening is presented. Feature selection is performed using fewer numbers of electrodes, which shows that frontal and parietal lobes provide more discriminative information to recognize human emotion. In [8], four different emotions in response to audio music using commercially available headset are recognized by using time, frequency, and wavelet domain features. The impact of different genres of music on human emotions has also been investigated.

Human emotion recognition in response to images, which involves the sense of vision, has also been investigated. In [32], two basic emotional states i.e., happiness and sadness is recognized by common spatial pattern and SVM using different frequency bands. International affective picture system emotion stimuli have been used to recognize four basic emotions based on human brain activity [25]. Features are extracted by using modified kernel estimation and artificial neural network is used to classify emotions. The brain response to pleasant and unpleasant touch has been investigated that engages the sense of tactile [40]. The sense of touch is evoked by caressing the human forearm with pleasant and unpleasant textile fabrics. The beta-power of EEG signals significantly distinguishes between pleasant and unpleasant touch. In [24], odor pleasantness classification is performed using EEG and ECG signals in response to ten different pleasant, unpleasant and neutral odors, which engages the olfactory sense. The frontal and central lobes of human brain are responsible for pleasantness classification, while ECG features proved to be less discriminative among the odors.

In [41], human emotions are recognized in response to videos by engaging two senses i.e., vision and auditory. Emotions are classified based on three classes of valence i.e., pleasant, neutral, and unpleasant and three classes of arousal i.e., calm, medium aroused, and activated based on EEG, pupillary diameter and gaze distance using SVM classifier. For EEG based classification, different frequency band features are used. A multimodal database is collected by using commercial physiological sensors like EEG, ECG, galvanic skin response (GSR), and facial activity, while viewing movie clips [42]. A relationship between human emotional states and their personality traits are also examined and results show that emotionally homogenous videos are better in revealing personality differences. Emotion assessment using EEG and peripheral signals in response to game at three difficulty levels is investigated [9]. It is established that different difficulty levels in a game evokes different emotions and playing the game multiple times with the same difficulty level gives raise to boredom. In [45], multidimensional information is extracted by decomposing EEG signals into intrinsic mode functions by using empirical mode decomposition. The extracted features are used to recognize human emotions using publicly available database for emotion analysis using physiological signals (DEAP) [23]. The frontal, pre-frontal, temporal, and parietal channels are found to be discriminative for emotion recognition. In [14], instead of using hand crafted features, deep learning is applied to extract features from raw EEG signals to recognize human emotions.

The aforementioned methods recognize human emotions using EEG signals in response to stimuli by engaging either vision, auditory, tactile, or olfaction sense individually, or combining both vision and auditory senses. Although, human emotion recognition in response to videos has been extensively studied, but an enriched multimedia content engaging more than two human senses is yet to be explored. In this paper, human emotion recognition using EEG signals in response to TEM that engages three human senses i.e., vision, auditory, and tactile is presented for the very first time to the best of our knowledge.

### 3 Tactile enhanced multimedia (TEM)

In this paper, TEM content is generated, which incorporates cold air and hot air effects by synchronizing a fan and a heater in traditional multimedia clips respectively. For this purpose, two commercial movies i.e., *'Tangled'* and *'Lord of the Rings'* are selected, which have a considerable number of views. A scene from each movie is selected that involves



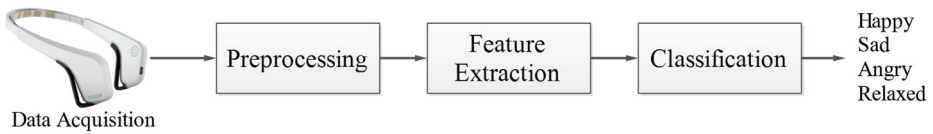
**Fig. 1** Screen grab from the scenes used in this study, Top row from the movie *‘Tangled’*, bottom row from the movie *‘Lord of the Rings’* synchronized with the effect of cold and hot air respectively

the possibility of enticing the tactile sense. The first multimedia clip is selected from a scene of the movie *‘Tangled’* having duration of 58 s. In this scene, a character runs on snow with the opposing effect of wind, which is identified on the basis of the unfurling of the character’s hair showing the presence of cold air. The cold air effect is added by synchronizing a fan with the multimedia content of the scene to generate TEM clip 1. The second multimedia clip of a duration 21 s is selected from the movie *‘Lord of the Rings’*, which shows a character moving around volcanoes. The scene starts normally and zooms after a few seconds showing an effect of hot air. This is evident by the sweat shown on the character’s face and unfurling of his hair. The hot air effect is added by identifying the timestamp to synchronize a heater with the multimedia content to generate TEM clip 2. The starting frame, ending frame and the synchronization frame of cold and hot air effect in TEM clip 1 and TEM clip 2 are shown in Fig. 1.

The starting frame for synchronization of fan and heater, and time duration of tactile effect for each TEM clip is identified by subjective evaluation. Different subjects rated the synchronization of cold and hot air with the scene by using like and dislike scores. Based on the feedback, the best synchronization point is identified and timestamps are noted. In order



**Fig. 2** The setup established for a user to experience TEM



**Fig. 3** The proposed methodology of human emotion recognition in response to TEM and traditional multimedia content using EEG signals

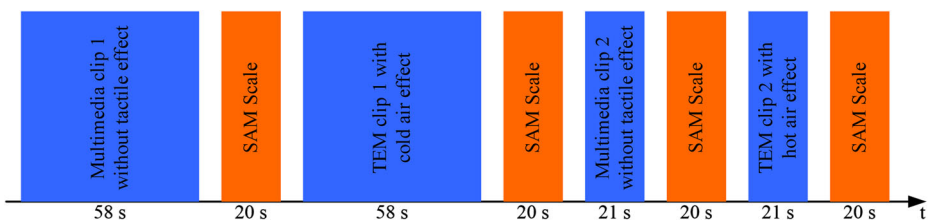
to generate TEM clips, a setup is established using an electric heater for hot air and a DC fan for cold air effect. These devices are controlled by an ARDUINO based controller, to turn the fan/heater on and off in synchronization with the timestamps in the clips. The setup established for experiencing TEM clips is shown in Fig. 2. Since tactile effect is mostly felt on the face and hands, the fan and heater are placed at a chest height. The viewer is provided with a comfortable chair and a 55-inch LED display is used for the display.

## 4 The proposed methodology

The steps involved in this study to classify emotions in response to TEM and traditional multimedia content are shown in Fig. 3. The details of each step are discussed in the following subsections.

### 4.1 Data acquisition

The EEG data is acquired by using commercially available MUSE EEG headset having four electrodes placed at the frontal and temporal regions of the brain. The data is recorded in an isolated room at a sampling rate of 256 Hz. A total of 21 participants are involved in this study including 11 males and 10 females, with an age between 19–23 years. All participants belonged to the same ethnicity and educational background. The procedure followed for EEG data acquisition in this study is shown in Fig. 4, where EEG recordings are performed during the phases presented in blue color. At the start of experiment, multimedia clip 1 from the movie ‘*Tangled*’ without tactile effect is displayed and EEG data is recorded. After viewing the clip, a 9-point SAM scale is displayed on the screen for valence and arousal scores, which is filled by the user with the help of a mouse. The valence score shows the pleasant or unpleasant state, whereas arousal score shows the calm or excited state of an individual. TEM clip 1 of the same scene from the movie ‘*Tangled*’, adding the effect of cold air is displayed and EEG data is recorded. The valence and arousal scores are gathered in response to TEM clip 1 with cold air effect. The multimedia clip 2 from the movie ‘*Lord*



**Fig. 4** Experimental procedure used for EEG data acquisition

of the Rings’ without adding the tactile effect is displayed and EEG data is recorded. After viewing the clip, a 9-point SAM scale is displayed on the screen for valence and arousal scores. TEM clip 2 of the same scene from the movie ‘Lord of the Rings’, adding the effect of hot air is displayed and EEG data is recorded. The valence and arousal scores are gathered in response to TEM clip 2 by using the SAM scale.

### 4.2 Preprocessing

The acquired EEG data is preprocessed to remove noise. Active noise suppression is performed using an on-board driven right leg (DRL-REF) feedback algorithm. The DRL circuits ensure a proper skin contact of the electrodes. Most of the information in EEG data lies below 30 Hz [35], hence a low-pass filter is used to remove high frequencies. Electrooculogram (EoG) and electromyogram (EMG) artifacts are kept at a minimum by instructing the participants to minimize their eye and muscle movements during data recording.

### 4.3 Feature extraction

Twelve time domain features are extracted from the preprocessed EEG signal  $s[i]$ , which includes absolute amplitude, absolute latency to amplitude ratio, average, mean of absolute values of first difference, mean of absolute values of second difference, mean of absolute values of first difference of normalized signal, mean of absolute values of second difference

**Table 1** Time domain features and their mathematical representation used for human emotion recognition

Feature	Mathematical representation
Absolute amplitude (AAMP)	$AAMP =  s_{max} $ , where $s_{max} = \max(s[i])$
Absolute latency to amplitude ratio (ALAR)	$ALAR =   \frac{t_{smax}}{s_{max}}  $ , where $t_{smax}$ is latency
Average ( $\mu_s$ )	$\mu_s = \frac{1}{N} \sum_{i=1}^N s[i]$ , where $N$ is the total number of samples
Mean of absolute values of first difference ( $\delta_s$ )	$\delta_s = \frac{1}{N-1} \sum_{i=1}^{N-1}  s[i + 1] - s[i] $
Mean of absolute values of second difference ( $\gamma_s$ )	$\gamma_s = \frac{1}{N-2} \sum_{i=1}^{N-2}  s[i + 2] - s[i] $
Mean of absolute values of first difference of normalized signal ( $\delta_s^n$ )	$\delta_s^n = \frac{1}{N} \sum_{i=1}^{N-1}  s^b[i + 1] - s^b[i] $ , where $s^b = \frac{s[i] - \mu_s}{\sigma_s}$ , $\mu_s$ and $\sigma_s$ are the mean and standard deviation of the signal $s[i]$
Mean of absolute values of second difference of normalized signal ( $\gamma_s^n$ )	$\gamma_s^n = \sum_{i=1}^{N-2}  s^b[i + 2] - s^b[i] $
Power ( $Pow$ )	$Pow = \frac{1}{N} \sum_{i=1}^N  s[i] ^2$
Energy ( $E$ )	$E = \sum_{i=1}^N  s[i] ^2$
Variance ( $\sigma_s^2$ )	$\sigma_s^2 = E[(s[i] - \mu_s)^2]$
Kurtosis ( $K$ )	$K = \frac{m_4}{\sigma_s^4}$ , where $m_4$ is the fourth moment of the signal $s[i]$
Skewness ( $S$ )	$S = \frac{m_3}{\sigma_s^3}$ , where $m_3$ is the third moment of the signal $s[i]$

of normalized signal, power, energy, variance, kurtosis, and skewness. The mathematical representation of these features is given in Table 1.

### 4.4 Classification

In this work, a support vector machine is used to classify four different emotions in response to TEM and traditional multimedia content based on extracted features from the pre-processed EEG signals. SVM has been widely used to classify emotions in response to different stimulus using EEG [2]. SVM minimizes upper boundary of the generalization error and is based upon the theory of structural risk minimization principle. SVM can be used for both classification and regression and is powerful because it can classify both linear and non-linear datasets by using kernel functions [37]. It transmits the data that is difficult to classify linearly to higher dimensional planes. A polynomial kernel based algorithm in SVM is used in this study and is defined as,

$$K(x, y) = (x^T y + c)^d, \tag{1}$$

where  $x$  and  $y$  are vectors in the input space and  $d$  is the degree of the polynomial.

### 5 Experimental results

In this study, each subject viewed two multimedia clips without tactile effects and the same two multimedia clips with cold and hot air tactile effect. After viewing each multimedia clip and its TEM version, a subject had to fill the valence and arousal score on the SAM scale. The valence and arousal values are used to check, whether the multimedia content with and without tactile effects is significantly different in terms of valence and arousal scores. Moreover, these valence and arousal scores are used to label the recorded EEG data with one emotion from each quadrant of the SAM scale i.e., happy, angry, sad, and relaxed [16].

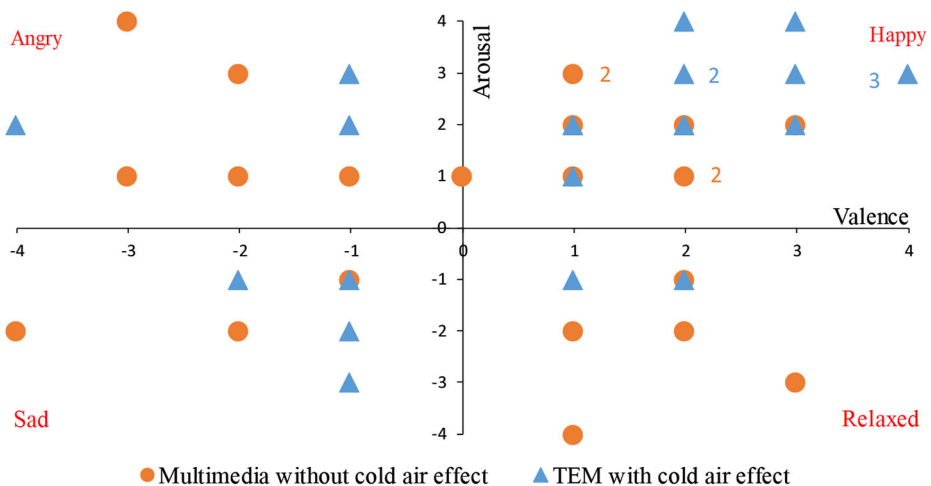


Fig. 5 Valence and arousal scores for multimedia clip from the movie ‘Tangled’ without tactile effect and its TEM version with cold air effect



The labeled EEG data are used for training the SVM classifier for emotion recognition in response to TEM and traditional multimedia content.

Valence and arousal value of all users in case of traditional multimedia and TEM with cold air effect is shown in Fig. 5. Similarly, valence and arousal values of all users in case of traditional multimedia and TEM with hot air effect is shown in Fig. 6. The user data points in the valence-arousal plane for traditional multimedia and TEM content are represented by orange and blue colors respectively. The numbers shown on the valence-arousal data points in orange and blue color represents the number of users who scored traditional multimedia or TEM clip at that point respectively. For positive valence and arousal values, data is labeled as happy. Similarly, for negative valence and arousal scores, data is labeled as sad. For negative valence and positive arousal, data is labeled as angry. Similarly, for positive valence and negative arousal, data is labeled as relaxed. It is evident, that TEM content emphasizes the valence and arousal scores as compared to traditional multimedia content due to the involvement of tactile sense in addition to auditory and vision sense. The experimental results are presented in two stages: (i) a statistical analysis of multimedia content with and without tactile effects, and (ii) a performance comparison of human emotion recognition using EEG signals in response to TEM and traditional multimedia content.

### 5.1 Statistical analysis

For a better understanding of the valence and arousal related discrimination in response to TEM and traditional multimedia content, a t-test is performed. The test decision is based on p-value and is considered significantly different if the value is less than 0.05. The t-test is applied on traditional multimedia content and cold air enhanced multimedia content for valence and arousal values. The p-value is 0.120 for valence and 0.001 for arousal. Similarly, a t-test is applied on the traditional multimedia content and TEM with hot air effect. The p-value is 0.0009 and 0.004 for valence and arousal values respectively. These p-values indicate that there is a significant difference between the valence and arousal scores in response to TEM and traditional multimedia content, which means traditional and TEM

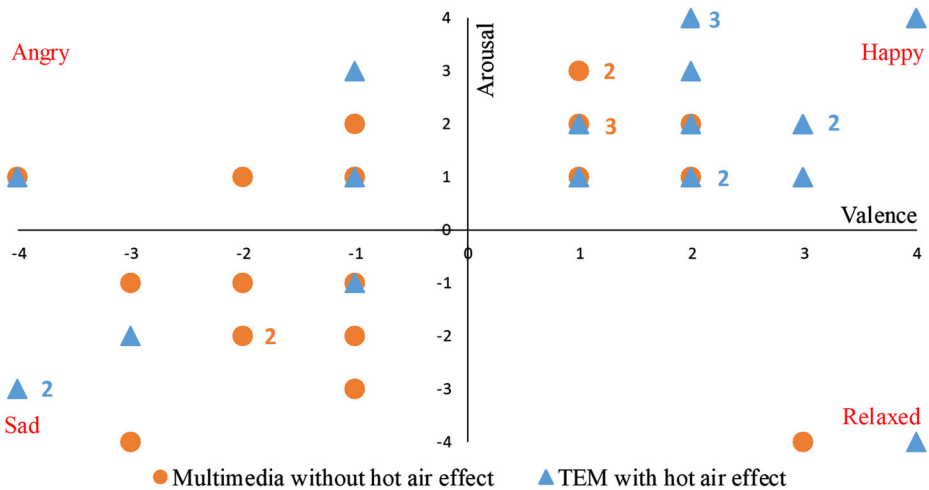


Fig. 6 Valence and arousal scores for multimedia clip from the movie 'Lord of the Rings' without tactile effect and its TEM version with hot air effect

content evokes different emotions. Arousal class is found significantly different for cold air enhanced multimedia content, whereas for hot air enhanced multimedia content both valence and arousal have p-value less than 0.05 representing significantly discriminating stimuli.

## 5.2 Performance comparison

The performance of the proposed emotion recognition in response to TEM and traditional multimedia using EEG signals is measured in terms of accuracy, mean absolute error (MAE), root mean squared error (RMSE), root absolute error (RAE), root relative squared error (RRSE), and Cohen's kappa values, which are calculated as,

$$MAE = \frac{1}{N} \sum_{i=1}^N |u_i - v_i|, \quad (2)$$

where,  $u_i$  and  $v_i$  are the coordinates of point  $i$ ,

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N |u_i - v_i|^2}, \quad (3)$$

$$RAE = \frac{\sum_{i=1}^N |u_i - v_i|}{\sum_{i=1}^N |\tilde{u}_i - v_i|}, \quad (4)$$

where  $\tilde{u}_i$  is the mean of  $u_i$ , and

$$RRSE = \sqrt{\frac{\sum_{i=1}^N (u_i - v_i)^2}{\sum_{i=1}^N (\tilde{u}_i - v_i)^2}}. \quad (5)$$

Cohen's kappa is used to evaluate the classifier performance for unbalanced classes [10]. The kappa value is calculated as,

$$k = \frac{p_r - p_e}{1 - p_e}, \quad (6)$$

where  $p_a$  is the relative observed agreement among the ground truth and the test data and  $p_e$  is the hypothetical probability of chance. The value of  $k$  equals 1, when ground truth and test data are in complete agreement, whereas a zero value of  $k$  means that the test data has no agreement with the ground truth data. A 10-fold cross validation is applied to recognize emotions using EEG signals in response to TEM and traditional multimedia. In 10-fold cross validation, all instances are divided into ten equal parts. Each part is once used as testing data, whereas the remaining nine parts are used as training data. The accuracies are averaged out to find the final accuracy of the classification algorithm.

The human emotion classification results are shown in Table 2 in terms of average accuracy, MAE, RMSE, RAE, RRSE, and kappa. Emotion classification against TEM content has the highest accuracy of 63.41%, whereas four emotions against traditional multimedia

**Table 2** Performance comparison of human emotion recognition in terms of accuracy, MAE, RMSE, RAE, RRSE and kappa in response to traditional multimedia and TEM content

Stimulus	Accuracy(%)	MAE	RMSE	RAE	RRSE	Kappa
Traditional multimedia	43.90	0.347	0.441	95.10	103.10	0.180
Tactile enhanced multimedia	63.41	0.306	0.392	111.67	107.30	0.150

content are classified at an accuracy of 43.90%. In terms of absolute and squared error rate, TEM has less absolute and relative error rate in comparison to the traditional multimedia content. For unbalanced classes, Cohen’s kappa value is used to compare the classification results. A value of  $k = 0.150$  for TEM content and  $k = 0.180$  for traditional multimedia content shows that the classifier gives better results for TEM content.

The performance of the proposed emotion recognition method is also evaluated in terms of precision (P), recall (R), and F-measure (F). Precision is the positive predicted value and recall is the sensitivity of the test data. Precision and recall are calculated using true positive (TP), false positive (FP), and false negative (FN) rates as follows,

$$P = \frac{TP}{TP + FP}, \tag{7}$$

$$R = \frac{TP}{TP + FN}. \tag{8}$$

Both precision and recall are a measure of relevance, whereas F-measure is the weighted harmonic mean of the precision and recall values and is given as,

$$F = 2 \times \frac{P \times R}{P + R}. \tag{9}$$

Precision, recall, and F-measure values against TEM and traditional multimedia content are shown in Fig. 7. It is evident that the proposed emotion recognition method using EEG signals gives higher precision, recall and F-measure values in response to TEM as compared to traditional multimedia content.

Recently, a number of methods have been proposed for emotion recognition using EEG [2]. Different methods report variant accuracies depending on the number of output classes and the number of EEG electrodes used. A comparative summary of the recently reported studies is shown in Table 3. These methods are compared in terms of type of multimedia content, features, number of emotions, number of electrodes, stimuli, number of subjects,

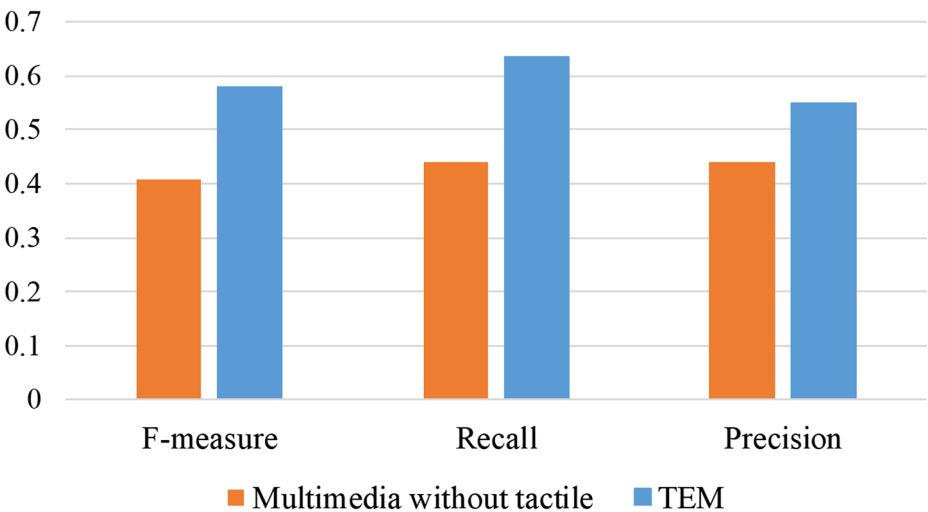


Fig. 7 Precision, recall, and F-measure against traditional multimedia and TEM content

**Table 3** Performance comparison of the proposed emotion recognition system using EEG with state-of-the-art methods

Article	Type	Features	No. of electrodes	Emotions	Multimedia type	No. of users (F/M)	Accuracy
[18]	User dependent	Frequency domain	1	Fear, relaxation	Traditional	8 (2/6)	92%
[28]	User dependent	Frequency domain	60	Positive, negative, neutral	Traditional	8 (5/3)	73%
[29]	User independent	Statistical	N/A	Joy, sadness, anger, surprise, disgust, fear, neutral	Traditional	9 (-/-)	36%
Proposed	User independent	Time domain	4	Happy, angry, sad, relaxed	Traditional TEM	21 (10/11)	43.90% 63.41%

and accuracy. In [18, 28, 29], emotions are recognized in response to videos using SVM as a classifier, where the number of emotions classified is 2, 3, and 7 respectively. In [18, 28], frequency domain features are used and techniques are user dependent, while in [29], statistical features are used and the technique is user-independent. Accuracies and number of electrodes are 92%(1), 73%(60) and 36%(N/A) respectively. The results of the proposed user-independent emotion recognition approach using EEG signals for four emotions in response to traditional multimedia content gives an accuracy of 43.90%. The proposed framework uses four electrodes and time domain statistical features, and the results are comparable with those reported in recent studies. Although, the methods presented in Table 3 for comparison have used different emotions, multimedia content, and EEG acquisition systems, but this comparison shows that the proposed technique has significant accuracy in classifying multiple emotions. The comparison also shows that the classification accuracy can be significantly improved by introducing a sense of tactile in multimedia content.

## 6 Conclusion

In this paper, emotion recognition using EEG in response to tactile enhanced multimedia content engaging more than two human senses is presented. Firstly, two TEM clips with cold and hot air effects are successfully generated by synchronization of a DC fan and a heater in accordance with accurate timestamps in the movie scenes. Traditional multimedia clips and their TEM versions were shown to 21 subjects and their EEG data was recorded along with the valence and arousal scores on the SAM scale. A t-test was performed on TEM and traditional multimedia content and showed a significant difference in terms of valence and arousal scores. This shows that TEM content has emphasized human emotions. Four different emotions were classified in response to TEM and traditional multimedia content using EEG signals. It is evident from classification results that emotions are recognized accurately in response to TEM content as compared to the traditional multimedia content. In future, we intend to recognize human emotion using physiological signals in response to olfaction and tactile enhanced multimedia content.

## References

1. Aftanas L, Reva N, Varlamov A, Pavlov S, Makhnev V (2004) Analysis of evoked eeg synchronization and desynchronization in conditions of emotional activation in humans: Temporal and topographic characteristics. *Neurosci Behav Physiol* 34(8):859–867
2. Alarcao SM, Fonseca MJ Emotions recognition using eeg signals: a survey. *IEEE Trans Affect Comput*. <https://doi.org/10.1109/TAFFC.2017.2714671>
3. Allen JJ, Coan JA, Nazarian M (2004) Issues and assumptions on the road from raw signals to metrics of frontal eeg asymmetry in emotion. *Biol Psychol* 67(1):183–218
4. Anagnostopoulos C-N, Iliou T, Giannoukos I (2015) Features and classifiers for emotion recognition from speech: a survey from 2000 to 2011. *Artif Intell Rev* 43(2):155–177
5. Balconi M, Lucchiari C (2008) Consciousness and arousal effects on emotional face processing as revealed by brain oscillations. A gamma band analysis. *Int J Psychophysiol* 67(1):41–46
6. Basar E, Basar-Eroglu C, Karakas S, Schurmann M (1999) Oscillatory brain theory: a new trend in neuroscience. *IEEE Eng Med Biol Mag* 18(3):56–66
7. Bethel CL, Salomon K, Murphy RR, Burke JL (2007) Survey of psychophysiology measurements applied to human-robot interaction. In: *The 16th IEEE international symposium on robot and human interactive communication, 2007. RO-MAN 2007. IEEE*, pp 732–737
8. Bhatti AM, Majid M, Anwar SM, Khan B (2016) Human emotion recognition and analysis in response to audio music using brain signals. *Comput Hum Behav* 65:267–275
9. Chanel G, Rebetez C, Bétrancourt M, Pun T (2011) Emotion assessment from physiological signals for adaptation of game difficulty. *IEEE Trans Syst Man Cybern Part A Syst Hum* 41(6):1052–1063
10. Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20(1):37–46
11. de Gelder B, De Borst A, Watson R (2015) The perception of emotion in body expressions. *Wiley Interdiscip Rev Cogn Sci* 6(2):149–158
12. Du Y, Zhang F, Wang Y, Bi T, Qiu J (2016) Perceptual learning of facial expressions. *Vision Res* 128:19–29
13. Fasel B, Luetttin J (2003) Automatic facial expression analysis: a survey. *Pattern Recogn* 36(1):259–275
14. Gao Y, Lee HJ, Mehmood RM (2015) Deep learning of eeg signals for emotion recognition. In: *2015 IEEE international conference on multimedia & expo workshops (ICMEW)*. IEEE, pp 1–5
15. Ghinea G, Timmerer C, Lin W, Gulliver SR (2014) Mulsemedia: state of the art, perspectives, and challenges. *ACM Trans Multimed Comput Commun Appl (TOMM)* 11 (1s) 17:1–23
16. Graziotin D, Wang X, Abrahamsson P (2015) Understanding the affect of developers: theoretical background and guidelines for psychoempirical software engineering. In: *Proceedings of the 7th international workshop on social software engineering*. ACM, pp 25–32
17. Heller W (1993) Neuropsychological mechanisms of individual differences in emotion, personality, and arousal. *Neuropsychology* 7(4):476–489
18. Jalilifard A, Pizzolato EB, Islam MK (2016) Emotion classification using single-channel scalp-eeg recording. In: *2016 IEEE 38th annual international conference of the engineering in medicine and biology society (EMBC)*. IEEE, pp 845–849
19. Kalsum T, Anwar SM, Majid M, Khan B, Ali SM (2018) Emotion recognition from facial expressions using hybrid feature descriptors. *IET Image Process* 12(6):1004–1012
20. Kim J, André E (2006) Emotion recognition using physiological and speech signal in short-term observation. In: *International tutorial and research workshop on perception and interactive technologies for speech-based systems*. Springer, pp 53–64
21. Kim J, André E (2008) Emotion recognition based on physiological changes in music listening. *IEEE Trans Pattern Anal Mach Intell* 30(12):2067–2083
22. Kim KH, Bang SW, Kim SR (2004) Emotion recognition system using short-term monitoring of physiological signals. *Med Biol Eng Comput* 42(3):419–427
23. Koelstra S, Muhl C, Soleymani M, Lee J-S, Yazdani A, Ebrahimi T, Pun T, Nijholt A, Patras I (2012) Deap: a database for emotion analysis; using physiological signals. *IEEE Trans Affect Comput* 3(1):18–31
24. Kroupi E, Vesin J-M, Ebrahimi T (2016) Subject-independent odor pleasantness classification using brain and peripheral signals. *IEEE Trans Affect Comput* 7(4):422–434
25. Lahane P, Sangaiah AK (2015) An approach to eeg based emotion recognition and classification using kernel density estimation. *Procedia Computer Science* 48:574–581
26. Lin H, Schulz C, Straube T (2016) Effects of expectation congruency on event-related potentials (erps) to facial expressions depend on cognitive load during the expectation phase. *Biol Psychol* 120:126–136
27. Lin Y-P, Wang C-H, Jung T-P, Wu T-L, Jeng S-K, Duann JR, Chen JH (2010) Eeg-based emotion recognition in music listening. *IEEE Trans Biomed Eng* 57(7):1798–1806

28. Liu S, Tong J, Xu M, Yang J, Qi H, Ming D (2016) Improve the generalization of emotional classifiers across time by using training samples from different days. In: 2016 IEEE 38th annual international conference of the engineering in medicine and biology society (EMBC). IEEE, pp 841–844
29. Matlovič T (2016) Emotion detection using eeg device. In: Information and informatics technologies student research conference (IIT.SRC), pp 1–6
30. Murray N, Qiao Y, Lee B, Karunakar A, Muntean G-M (2013) Subjective evaluation of olfactory and visual media synchronization. In: Proceedings of the 4th ACM multimedia systems conference. ACM, pp 162–171
31. Murray N, Qiao Y, Lee B, Muntean G-M (2014) User-profile-based perceived olfactory and visual media synchronization. *ACM Trans Multimed Comput Commun Appl (TOMM)* 10 (1s) 11:1–24
32. Pan J, Li Y, Wang J (2016) An eeg-based brain-computer interface for emotion recognition. In: 2016 international joint conference on neural networks (IJCNN). IEEE, pp 2063–2067
33. Picard RW, Vyzas E, Healey J (2001) Toward machine emotional intelligence: analysis of affective physiological state. *IEEE Trans Pattern Anal Mach Intell* 23(10):1175–1191
34. Qayyum H, Majid M, Anwar SM, Khan B (2017) Facial expression recognition using stationary wavelet transform features. *Math Probl Eng* 2017:9
35. Sanei S, Chambers JA (2007) EEG Signal processing. Wiley Online Library
36. Sarlo M, Buodo G, Poli S, Palomba D (2005) Changes in eeg alpha power to different disgust elicitors: the specificity of mutilations. *Neurosci Lett* 382(3):291–296
37. Şen B, Peker M, Çavuşoğlu A, Çelebi FV (2014) A comparative study on classification of sleep stage based on eeg signals using feature selection and classification algorithms. *J Med Syst* 38(18):1–21
38. Schmidt LA, Trainor LJ (2001) Frontal brain electrical activity (eeg) distinguishes valence and intensity of musical emotions. *Cognit Emot* 15(4):487–500
39. Schutter DJ, Putman P, Hermans E, van Honk J (2001) Parietal electroencephalogram beta asymmetry and selective attention to angry facial expressions in healthy human subjects. *Neurosci Lett* 314(1):13–16
40. Singh H, Bauer M, Chowanski W, Sui Y, Atkinson D, Baurley S, Fry M, Evans J, Bianchi-Berthouze N (2014) The brain's response to pleasant touch: an eeg investigation of tactile caressing. *Front Hum Neurosci* 8:893
41. Soleymani M, Pantic M, Pun T (2012) Multimodal emotion recognition in response to videos. *IEEE Trans Affect Comput* 3(2):211–223
42. Subramanian R, Wache J, Abadi MK, Vieriu RL, Winkler S, Sebe N (2018) Ascertain: emotion and personality recognition using commercial sensors. *IEEE Trans Affect Comput* 9(2):147–160
43. Sulema Y (2016) Mulsemedia vs. multimedia: state of the art and future trends. In: 2016 international conference on systems, signals and image processing (IWSSIP). IEEE, pp 1–5
44. Yuan Z, Chen S, Ghinea G, Muntean G-M (2014) User quality of experience of mulsemedia applications. *ACM Trans Multimed Comput Commun Appl (TOMM)* 11 15(1s):1–19
45. Zhuang N, Zeng Y, Tong L, Zhang C, Zhang H, Yan B (2017) Emotion recognition from eeg signals using multidimensional information in emd domain. *Biomed Res Int* 2017:9



**Aasim Raheel** received his BSc. and MSc. in Computer Engineering from University of Engineering and Technology (UET) Taxila in 2010 and 2015 respectively. Currently he is pursuing his PhD in Computer Engineering from UET Taxila. His research interest includes affective computing, facial expression and emotion recognition using EEG, and brain computer interface.



**Syed Muhammad Anwar** received BSc. in Computer Engineering with honors from University of Engineering and Technology (UET) Taxila, Pakistan in 2005, MSc. in Data Communications with distinction and PhD in Electronic and Electrical Engineering from the University of Sheffield, UK in 2007 and 2012 respectively. He is currently an Associate Professor at Department of Software Engineering, UET Taxila. His research interests include medical imaging, data communication, and human computer interaction.



**Muhammad Majid** received BSc. in Computer Engineering with honors from University of Engineering and Technology (UET) Taxila, Pakistan in 2005, MSc. in Data Communications with distinction and PhD in Electronic and Electrical Engineering from the University of Sheffield, UK in 2007 and 2011 respectively. He is currently an Associate Professor at Department of Computer Engineering, UET Taxila and heading Signal, Image, Multimedia Processing and LEarning (SIMPLE) research group since January 2012. He was a recipient of Faculty Development Program Scholarship of Higher Education Commission (HEC) Pakistan for his MSc and PhD Studies. His research interests include image and video coding, multimedia signal processing, biomedical signal processing and emotion classification in response to multimedia content.