International Journal of Enhanced Research in Science Technology & Engineering, ISSN: 2319-7463
Vol. 3 Issue 4, April-2014, pp: (230-235), Impact Factor: 1.252, Available online at: www.erpublications.com

# A Survey: Emotion Detection via Facial Expression

Shweta Shukla[1], Ashish Shukla[2], Vishal Singh[3]
[123]M.Tech., Department of Computer Science and Engineering,
Pranveer Singh Institute of Technology, Kanpur, India

**Abstract: Emotion detection has been an active research topic for quite a few years and has wide applications in various domains. Emotion detection can convey the emotional state, intention, psychology or psychopathology of a person. During interactions, they can convey certain non-behavioral cues and hence proper recognition of emotion can improve the quality of interaction. Human beings can very well identify emotion but it is still quite a challenge for a machine to do so. Moreover, they form an important component of human- machine interaction and hence correct recognition of subtle expressions would be indispensable. Emotional states significantly affect the cognition and behaviors of human. Scientific findings suggest an increasingly large number of important functions of emotion [1]. Emotion recognition based on information technology is an important research topic in the field of neural engineering. Principal Component Analysis (PCA), Independent Component Analysis (ICA), Active Appearance Model (AAM) based emotion recognition has been done in the last decade. The AAM fitting algorithm was first developed by T.F. Coots, G.J. Edwards and C.J. Taylor [7]. The AAM fitting is an algorithm for matching a statistical model of object shape and appearance to a new image. The algorithm uses the difference between the shape vector and the appearance of the current estimate and the target image to run an optimization process. By taking advantage of the least squares techniques, it can match to new images very efficiently. The human visual system can recognize special variations up to a certain extent. Motion Magnification amplifies these subtle spatial variations to reveal certain hidden information. We have implemented Motion Magnification which converts subtle facial expressions into exaggerated ones and hence classify them accurately.**

**Keywords: Emotion Detection, Emotion Recognition, Active Appearance Model.**

## Introduction

There are different emotion classification systems. The taxonomy can be seen from two perspectives: dimensional and discrete one [2]. Plutchik defines eight basic emotion states: anger, fear, sadness, disgust, surprise, anticipation, acceptance and joy. All other emotions can be formed by these basic ones, for example, disappointment is composed of surprise and sadness [3]. Another approach towards emotion classification is advocated by Paul Ekman. He revealed the relationship between facial expressions and emotions. In his theory, there are six emotions associated with facial expressions: anger, disgust, fear, happiness, sadness, and surprise. Later he expanded the basic emotion by adding: amusement, contempt, contentment, embarrassment, excitement, guilt, and pride in achievement, relief, satisfaction, sensory pleasure, and shame.

From the dimensional perspective, the most widely used one is the bipolar model where arousal and valence dimensions are considered. This emotion classification approach is advocated by Russell [5]. Here, the arousal dimension ranges from not aroused to excited, and the valence dimension ranges from negative to positive. Another fundamental dimension is an approach-withdrew dimension which is based on the motivating aspects of the emotion [2]. For example, in this theory, anger is an approach motivated emotion in some cases, as it could encourage the person to make effort to change the situation. The dimensional model is preferable in emotion recognition experiments due to the following advantage: dimensional model can locate discrete emotions in its space, even when no particular label can be used to define a certain feeling.

Emotions are the key to the human decision making processes since decisions and actions are primary irrational and not cognitive.

Recent research shows that the emotional sphere is much quicker than the rational one. A few milliseconds after we come across an object (a person, an object, a word, a picture), a primary emotional response occurs. The psychologist Paul Ekman performed a series of videotaped experiments in which he followed the minor changes in a person's facial structure when coming across these primal interactions. These experiments showed great emotional responses within the first second of the interaction. Within a few milliseconds changes occurred in movements of facial muscles. Within less than half a second a significant emotional reaction had occurred (i.e. anger, fear, pleasure, sadness). Physiological changes such as an increased heartbeat occurred within fractions of a second. Another research study, performed by John A. Bargh in 1994, dealt with what occurs in the subconscious in the first second. This research showed that within the first milliseconds we not only understand what we see and feel but also decide whether we like it or not. Carol Kinsey Gorman, reported at Forbes about researchers from NYU who found that it takes just 7 seconds to make a first impression.

## Literature Review

### A. EEG-based emotion perception during music listening[8]

Proceeding of the 12th international conference on music perception and cognition and the 8th triennial conference of the European society for the cognition science of music, July 23-28, 2013.

**Explanation:** In this paper, correlations between electroencephalographic (EEG) activity and emotional responses during music listening were investigated. EEG activity was recorded in different regions without a-priori defining regions of interest. The analysis of the data was performed in both alpha and theta bands. The results in alpha band confirm the hemispheric specialization hypothesis for emotional valence. Positively valence emotions (happy and serene) elicited greater relative left EEG activity, whereas negatively valence emotions (angry and sad) elicited greater relative right EEG activity. The results show interesting findings related to the affective dimension (arousal and valence) by electrodes in different brain regions that might be useful in extracting effective features for emotion recognition applications. Moreover, theta asymmetries observed between pleasant and unpleasant musical excerpts support the hypothesis that theta power may have a more important role in emotion processing than previously believed.

### B. EEG-based discrimination of music appraisal judgments using ZAM time-frequency distribution [9]

Proceeding of the 12th international conference on music perception and cognition and the 8th triennial conference of the European society for the cognition science of music, July 23-28, 2013.

**Explanation:** In this paper, Although EEG-based emotional responses to music have Been studied extensively, little research has been conducted for the discrimination of music appraisal judgments. In the psychology-based literature music appraisal is mainly interpreted in terms of affective experiences, such as emotional resonance, aesthetic awe and thrill (Konečni, 2005; Schubert, 2007; Evans & Schubert, 2008). In the field of neurophysiology, EEG-based evidence of brain activation due to emotion-evocative music is mainly reported. Altenmüller, Schurmann, Lim, and Parlitz (2002) found that pleasant music causes left frontal brain activation, while unpleasant music causes right and slightly bilateral frontal activation. Similar evidence has been produced by the study of Schmidt and Trainor (2001). Another EEG-based study showed that consonant (pleasant) music causes greater midline brain activity compared to dissonant (unpleasant) music (Slammer, Grigutsch, Fritz, &Koelsch, 2007). As far as EEG-based emotion recognition is concerned, Lin et al. (2010) achieved an accuracy of 82.29 ± 3.06% for the classification of distinct emotions due to music listening, using spectrogram-based feature extraction and support vector machines (SVM). However, the aforementioned electrophysiological evidence arises from the listening of positively/negatively valence music that cannot be directly mapped to liked/disliked music. For instance, depending on the listeners' mood, sad-sounding musical excerpts can be occasionally preferred instead of happy-sounding ones. In general, evidence of brain activity related to emotional responses is reported in the majority of EEG frequency bands, i.e., theta, alpha, beta and gamma. Frontal midline theta power modulation is suggested to reflect affective processing during consonant/dissonant music (Slammer et al., 2007). The alpha-power asymmetry on the prefrontal cortex has been proposed as an index for the discrimination between positively and negatively valence emotions (Davidson, 2004). Moreover, beta activity has been associated with emotional arousal modulation (Aftanas, Reva, Savotina, &Makhnev, 2004), while activity in the gamma band is also related to arousal effects (Keil et al., 2001). In this framework, the present study aims at classifying listeners' EEG responses that relate to music liking or disliking judgments. Specifically, the main objectives of this work are: 1) to propose energy-based time-frequency (TF) features for an efficient classification and 2) to associate the EEG-based results with evidence from the existing literature on music evoked emotions and emotional responses in general.

### C. EEG-Based Emotion Recognition in Listening Music by Using Support Vector Machine and Linear Dynamic System [10]

T. Huang et al. (Eds.): ICONIP 2012, Part IV, LNCS 7666, pp. 468–475, 2012.

**Explanation:** This paper focuses on the variation of EEG at different emotional states. We use pure music segments as stimuli to evoke the exciting or relaxing emotions of subjects. EEG power spectrum is adopted to form features, power spectrum, differential asymmetry, and rational asymmetry. A linear dynamic system approach is applied to smooth the feature sequence. Minimal-redundancy-maximal-relevance algorithm and principal component analysis are used to reduce the dimension of features. We evaluate the performance of support vector machine, k-nearest neighbor classifiers and least-squares classifiers. The accuracy of our proposed method reaches 81.03% on average. And we show that the frequency bands, beta and theta, perform better than other frequency bands in the task of emotion recognition.

**D. Real-time EEG-based Emotion Recognition and its Applications [11]**

**Explanation**: Since emotions play an important role in the daily life of human beings,the need and importance of automatic emotion recognition has grown with increasing role of human computer interface applications. Emotion recognition could be done from the text, speech, facial expression or gesture. In this paper, we concentrate on recognition of "inner" emotions from electroencephalogram (EEG) signals. We propose real-time fractal dimension based algorithm of quantification of basic emotions using Arousal-Valence emotion model. Two emotion induction experiments with music stimuli and sound stimuli from International Affective Digitized Sounds (IADS) database were proposed and implemented. Finally, the real-time algorithm was proposed, implemented and tested to recognize six emotions such as fear, frustrated, sad, happy, pleasant and satisfied. Real-time applications were proposed and implemented in 3D virtual environments. The user emotions are recognized and visualized in real time on his/her avatar adding one more so-called "emotion dimension" to human computer interfaces. An EEG enabled music therapy site was proposed and implemented. The music played to the patients helps them deal with problems such as pain and depression. An EEG-based web-enable music player which can display the music according to the user's current emotion states was designed and implemented.

**E. Active Appearance Models for Face Recognition [12].**
Paul Ivan, April 4, 2007

**Explanation:** A growing number of applications are starting to use face recognition asthe initial step towards interpreting human actions, intention, and behavior, as a central part of next-generation smart environments. Recognition of facial expressions is an important example of face-recognition techniques used in these smart environments. In order to be able to recognize faces, there are some difficulties to overcome. Faces are highly variable, deformable objects, and can have very different appearances in images depending on pose, lighting, expression, and the identity of the person. Besides that, face images can have different backgrounds, differences in image resolution, contrast, brightness, sharpness, and color balance. This paper describes a model-based approach, called Active Appearance Models, for the interpretation of face images, capable of overcoming these difficulties. This method is capable of 'explaining' the appearance of a face in terms of a compact set of model parameters. Once derived, this model gives the opportunity for various applications to use it for further investigations of the modeled face (like characterize the pose, expression, or identity of a face). The second part of this paper describes some variations on Active Appearance Models aimed at increasing the performance and the computational speed of Active Appearance Models.

## Emotion Recognition Algorithm

There are so many algorithms for emotion recognition but we are using the improved Active Appearance model (AAM) for emotion recognition. Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar: The basic idea behind AAM: PCA (Principal component analysis) is used to find the mean shape and main variations of the training data to the mean shape. After finding the Shape Model, all training data objects are deformed to the main shape, and the pixels converted to vectors. Then PCA is used to find the mean appearance (intensities), and variances of the appearance in the training set. Both the Shape and Appearance Model are combined with PCA to one AAM-model. By displacing the parameters in the training set with a known amount, an model can be created which gives the optimal parameter update for a certain difference in model-intensities and normal image intensities. This model is used in the search stage.

**Working of AAM**

This algorithm allows us to find the parameters of the model, which generate a synthetic image as close as possible to a particular target image, assuming a reasonable starting approximation.
Interpretation of a previously unseen image is seen as an optimization problem in which the difference between this new image and the model (synthesized) image is minimized.

$\delta I = I_i - I_m$, (1)

Where, $I_i$ is the vector of grey-level values in the image and $I_m$is the vector of grey-level values for the current model parameters. We wish to minimize $\Delta = |\delta I|^2$, by varying the model parameters, c. This appears to be a difficult high-dimensional optimization problem, but in Cootes et al. pose that the optimal parameter update can be estimated from $\delta I$.

The spatial pattern in δI, encodes information about how the model parameters should be changed in order to achieve a better fit. There are basically two parts to the problem:

1. Learning the relationship between _I and the error in the model parameters δc.

2. Using this knowledge in an iterative algorithm for minimizing Δ.

The appearance model has one compact parameter vector c, which controls the shape and the texture (in the model frame) according to:

$x = \bar{x} + Q_s c$ $g = \bar{g} + Q_g c$ (2)

Where,

$Q_s = P_s W^{-1}_s$ $P_{cs}$, $Q_g = P_g P_{cg}$ (3)

Where $\bar{x}$ is the mean shape and $\bar{g}$ is the mean texture in a mean-shaped patch. A shape in the image frame, X, can be generated by applying a suitable transformation to the point, x : X = St(x). Valid transformations are, scaling (s), an in-plane rotation (ɵ), and a translation (lx, ly). If for linearity we represent the scaling and rotation as (sx, sy) where sx = (s cosɵ −1) and

4To find a reasonable starting position, often a separate module/application is used, which has a fast way of finding an estimate of the position of a face in an image.

### The Active Appearance Search Algorithm

sy = s sinɵ, then the pose parameter vector t = (sx, sy, lx, ly)T is zero for the identity transformation and St+δt(x) _ St(Sδt(x)). Now, in homogeneous co-ordinates, t corresponds to the transformation matrix :

$$S_t = \begin{pmatrix} 1+s_x & -s_y & l_x \\ s_y & 1+s_x & l_y \\ 0 & 0 & 1 \end{pmatrix} \quad (4)$$

For the AAM we must represent small changes in pose using a vector, δt. This is to allow us to predict small pose changes using a linear regression model of the form δt = Rg. For linearity the zero vector should indicate no change, and the pose change should be approximately linear in the vector parameters. This is satisfied by the above parameterization. The AAM algorithm requires us to find the pose parameters t0 of the transformation obtained by first applying the small change given by δt, then the pose transform given by t. Thus, find t0 so that St0(x) = St(Sδt(x)). Now it can be shown that for small changes, Sδt1(Sδt2(x)) δS(δt1+δt2)(x).

From the appearance model parameters c and shape transformation parameters, t we get the position of the model points in the image frame X. This gives the shape of the image patch to be represented by the model. During the matching phase we sample the pixels in this region of the image, gimage, and project into the texture model frame, gs = T−1u (gimage), with Tu. Then, the current model texture is given by gm = $\bar{g}$ +Qqc. The current difference between model and image in the normalized texture frame is then:

r(p) = gs–gm (5)

Where (p) are the parameters of the model, pT = (cT |tT |uT). A scalar measure of difference is the sum of squares of elements of r, E(p) = r(p)T r(p). A first order Taylor expansion of (5) gives,

$$r(p+\delta p) = r(p) + \frac{\partial r}{\partial p}\delta p, \quad (6)$$

Where the ijth element of matrix $\frac{\partial r}{\partial p}$ is $\frac{dr_i}{dp_j}$

Suppose during matching our current residual is r. We wish to choose δpso as to minimize |r(p+δp)|2. By equating (6) to zero we obtain the RMS (root mean squared) solution.

$$\delta p = -Rr(p), where R = (\frac{\partial r}{\partial p}^T \frac{\partial r}{\partial p})^{-1} \frac{\partial r}{\partial p}^T \quad (7)$$

Normally it would be necessary to recalculate $\frac{\partial r}{\partial p}$ at every step, an expensive operation. However, we assume that since it is being computed in a normalized reference frame, it can be considered approximately fixed. We can thus estimate it once from our training set. We estimate $\frac{\partial r}{\partial p}$ by numeric differentiation, systematically displacing each parameter from the known optimal value on typical images and computing an average over the training set. Residuals at displacements of differing magnitudes are measured (typically up to 0.5 standard deviations of each parameter) and combined with a Gaussian kernel to smooth them. We then precompute R and use it in all subsequent searches with the model. Now if we have computed the matrix R, we can construct an iterative method for solving the optimization problem. Given a current estimate of model parameters, c, the pose t, the texture transformation u, and the image sample at the current estimate gimage, one step of the iterative matching procedure is as follows:

1. Project the texture sample into the texture model frame using gs=T−1u (gimage).

2. Evaluate the error vector, r(p) = gs − gm, and the current error, E =|r(p)|2.

3. Compute the predicted displacements, δp = −Rr(p).

4. Update the model parameters ˆp ! p + kδp, where initially k = 1.

5. Calculate the new points, bX and the model frame texture ˆgm.

6. Sample the image at the new points to obtain ˆgimage.

7. Calculate a new error vector, r(ˆp) = T−1ˆu (ˆgimage) − ˆgm.

8. If |r(ˆp)|2 < E then accept the new estimate (record p = ˆp), otherwise try at k = 0.5, k = 0.25,etc.

9. Repeat this procedure until no improvement is made to the error, |r(p)|2, and convergence is declared.

## Problem Statement

An improved approach to emotion detection using modified AAM algorithm. In our algorithm we focused on main three constraints are Time, accuracy, number of recognized emotions. Because previous algorithms are not focused on all three constraints together so, our aim is to design an algorithm which focus on these basic and make the algorithm efficient than previous algorithms.

**Issues we will focus on:**

1. **Time constraint:** The performance time for the feature extraction and time of classification.

2. **Accuracy:** The accuracy of the emotion recognition still needs to be improved. The accuracy decreases when more emotions are needed to be recognized.

3. **Number of the recognized emotions:** Although there are varieties of emotional states to describe the human's feelings, until now only limited types of emotions can be recognized. But our algorithm recognized at least 5 emotions.

## Conclusion/Future Work

In this paper presents efficient Active Appearance Model for emotion recognition but there are another analysis such as Principal Component Analysis and Independent Component Analysis play an important role for emotion recognition. AAM algorithm is an improved algorithm for emotion recognition but there are three constraints Time, Accuracy, Number of the recognized emotions by which we can design a most efficient algorithm. As well as this paper will help for researchers who work on emotion detection via facial expression.

## References

[1]. Picard, R.W., Klein, J.: Toward Computers that Recognize and Respond to User Emotion:Theoretical and Practical Implications.Interacting with Computers14(2), 141–169 (2002).
[2]. Mauss, I.B., Robinson, M.D.: Measures of emotion: A review. Cognition and Emotion23(2), 209–237 (2009).

[3]. Plutchik, R.: Emotions and life : perspectives from psychology, biology, and evolution.American Psychological Association, Washington, DC, 1st edn. (2003).

[4]. Ekman, P.: Basic emotions. In: Dalgleish, T., Power, M. (eds.) Handbook of Cognition and Emotion.Wiley, New York (1999).

[5]. Russell, J.A.: Affective space is bipolar. Journal of Personality and Social Psychology 37(3),345–356 (1979).

[6]. Chanel, G.: Emotion assessment for affective-computing based on brain and peripheralsignals. Ph.D. thesis, University of Geneva, Geneva (2009).

[7]. C.J. Taylor T.F. Cootes, G.J. Edwards. Active appearance models. In Computer VisionECCV98, volume 1407, pages 484 { 498. Springer Berlin Heidelberg, 1998.

[8]. KonstantinosTrochidis and Emmanuel Bigand, EEG based emotion perception during listening music, proceeding of the 12th international conference on music perception and cognition and the 8th triennial conference of the European society for the cognition science ofmusic ,july 23-28,2013.

[9]. SteliosHadjidimitriouand LeontiosHadjileontiadis, EEG-based discrimination of music appraisal judgments using ZAM time-frequency distribution, proceeding of the 12th international conference on music perception and cognition and the 8th triennial conferenceof the European society for the cognition science of music , July 23-28,2013.

[10]. Ruo-Nan Duan1, Xiao-Wei Wang1, and Bao-Liang Lu, EEG-Based Emotion Recognition in Listening Music by Using Support Vector Machine and Linear Dynamic System ,T. Huanget al.(Eds.): ICONIP 2012, Part IV, LNCS 7666, pp. 468–475, 2012.

[11]. Yisi Liu, Olga Sourina, and Minh Khoa Nguyen, Real-time EEG-based Emotion Recognition and its Applications, {LIUY0053,EOSourina,RaymondKhoa}@ntu.edu.sg.

[12]. Paul Ivan, Active Appearance Models for Face Recognition, April 4, 2007.