AUBADE - A WEARABLE EMG AUGMENTATION SYSTEM FOR ROBUST EMOTIONAL UNDERSTANDING

C. D. Katsis *, N. Katertsidis**, G. Ganiatsas ** and D.I. Fotiadis**

* Department of Medical Physics, Medical School, University of Ioannina, Ioannina, Greece
** Unit of Medical Technology and Intelligent Information Systems, Department of Computer Science, University of Ioannina, Ioannina, Greece

me00526@cc.uoi.gr, nkaters@cc.uoi.gr, gganiats@cc.uoi.gr, fotiadis@cc.uoi.gr

Abstract: AUBADE is a wearable platform which monitors and recognises the emotional state of its users in real time, using selected biosignals. It is designed to be applied to persons operating under extreme stress conditions such as car racing drivers. It can also be used in a variety of health care applications mainly in the neurology and psychology field. AUBADE system's clinical applications are mainly based on the ability of supporting clinical diagnosis. Furthermore AUBADE proposes the implementation of a real-time 3D facial representation module which animates a generic face mesh, according to the user's facial muscle movements. The system has been preliminary validated using data obtained from four drivers in simulated race conditions. The emotional classes identified were high stress, low stress, disappointment, euphoria and neutral face. The average classification rate for the five emotional classes was 86%. The system pilot application will be further tested and evaluated on more subjects and under realistic conditions.

Introduction

The use of emotional intelligence in computers is a field of increasing importance. In many ways emotions are one of the last and least explored frontiers of intuitive human computer interaction. This may be explained by the fact that computers are traditionally viewed as logical and rational tools, something that is incompatible with the often irrational and seeming illogical nature of emotions [1]. It is also apparent that we as humans, while extremely good at feeling and expressing emotions, still cannot agree on how they should best be defined [2].

After a century of research, there is little agreement about a definition of emotions and many theories have been proposed. A number of these could not be verified until recently when improved measurement of specific physiological signals became available. In general emotions are short-term, whereas moods are long-term, and temperaments or personalities are very long-term [3]. Furthermore, the physiological muscle movements, comprising what looks to an outsider to be a facial expression, may not always correspond to a real underlying emotional state.

Emotion consists of more than its outward physical expression; it also consists of internal feelings and thoughts, as well as other internal processes of which the person experiencing the emotion may not be aware [4].

As machines and people begin to co-exist and cooperatively share a variety of tasks, the need for machines to constantly evaluate the affective condition of humans becomes more than apparent. This has prompted researchers in the engineering and computer science communities to develop automatic ways for computers to recognise emotions. The labelling of emotions into different states led most researchers to use pattern recognition approaches for recognising emotions, using different modalities as inputs to the emotion recognition models. The work in automatic understanding of affective condition has focused on classification of the universal expressions defined by Ekman [5]. These expressions are sadness, anger, fear, disgust, surprise, happiness, neutral and contempt. Thus, the implemented algorithms were tailored towards developing models to recognise the universal expressions from static images or video sequences [6-10]. These facial actions are essentially facial phonemes, which can be assembled to form facial expressions. There are also recent methods that employ a combination of audio and video signals for emotion recognition [11-17].

One of the hallmarks in emotion theory is whether distinct physiological patterns accompany each emotion [18]. Ekman et al. [19] and Winton et al. [20] provided some of the first findings showing significant differences in autonomic nervous system signals according to a small number of emotional categories or dimensions, but there was no exploration of automated classification. Flidlund and Izard [21] appear to be the first who applied pattern recognition (linear discriminants) on the classification of four different emotions (happiness, sadness, anger, fear) from physiological signals, attaining rates of 38-51 % accuracy. Similar efforts aimed at finding physiological correlates, focusing on t-tests or analysis of variance comparisons and combining data over many subjects, where each was measured for a relatively small amount of time [22-23]. Finally Picard et al. [4] classified physiological patterns for a set of eight emotions (including neutral) by applying pattern recognition

ISSN: 1727-1983 © 2005 IFMBE

techniques and focusing on felt emotions of a single subject over sessions spanning many weeks.

Although dealing with emotion recognition, the aforementioned techniques present the following limitations: (i) they are all materialized in laboratory environments therefore their effectiveness in real conditions is unknown, (ii) they are not real time and (iii) the data acquisition systems used for them are not wearable.

The work in this paper is novel, since it presents a system that automatically monitors and classifies the psychological condition of human subjects from a set of emotions. AUBADE estimates the emotional state of human subjects by classifying vectors of features extracted from: facial Electromyogram (EMG), Respiration, Electrodermal Activity (EDA) and Electrocardiogram (ECG).

Furthermore AUBADE proposes the implementation of a real time 3D facial representation module which animates a generic face mesh according to the user's facial muscle movements. The system is designed to be applicable to persons operating under extreme stress conditions, such as car racing drivers. Medical applications are mainly based on the ability of supporting clinical diagnosis related to all the pathologies according to which the patient's capability to feel and express emotions is limited or totally absent.

Materials and Methods

The AUBADE system consists of: (a) a multi-sensorial wearable, (b) a data acquisition and wireless communication module, (c) a feature extraction module, (d) a facial representation module which animates a generic face mesh according to the user muscle movements, (e) the AUBADE databases where the acquired signals, the facial animation videos and personal medical information are stored and (f) the intelligent emotion recognition module where vectors of features are used in order to determine subject's basic emotions.

The architecture of the AUBADE system is shown in Fig. 1. A more detailed description of the AUBADE system's functionalities and modules follows:

a) *The multi-sensorial wearable:* It is a non-invasive, ergonomic, comfortable and easy to use wearable that includes a number of bio-sensors gathering raw physiological data (facial EMG, respiration, EDR and ECG). The wearable is composed of three pieces: i) the balaclava containing sixteen EMG textile fireproof sensors, ii) the 3-lead ECG and respiration sensors on the thorax of the driver and iii) the EDR textile and fireproof sensor placed inside the drivers glove.

b) The data acquisition and wireless communication module: The signal acquisition unit consists of both hardware (Data Acquisition card) and software components. It appropriately collects, filters, pre-processes formats and stores all biosignals obtained from the sensors of the wearable. The pre-processing procedure (sampling rate and filters used) is presented in Table 1. The resolution used during signal digitization is 12 bit.

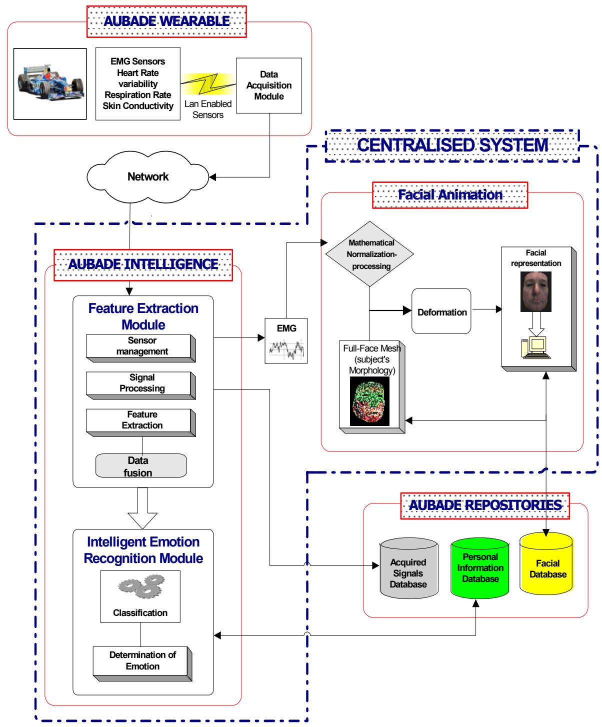

Figure 1: The AUBADE System Architecture

The communication module is responsible for the secure transfer of the biosignals - collected and processed by the Data Acquisition Unit - to the Intelligent Emotion recognition module for further analysis. The user measurements are transferred through either a fixed or a wireless LAN to the intelligent emotion recognition module for further analysis.

Table 1: Biosignals Preproseccing

Signal	Sampling Rate	Filters used
Facial EMG (16 channels)	1000 Hz	Low pass (500 Hz)
ECG	500 Hz	Low pass (100 Hz)
EDR	50 Hz	Moving average
respiration	50 Hz	Moving average

Moreover, appropriate security and privacy mechanisms have been implemented. Such mechanisms include authentication, secure protocols and strong encryption for wireless communication, protection of personal data and medical information over wireless networks.

(c) The feature extraction module: The pre-processed biosignals are converted into vectors of extracted features that can be used by the Intelligent Emotion Recognition module in order to determine subject's basic emotions. The selected features provide a combination of simple statistics and complicated characteristics which are related to the nature of the

physiological signals and the underlying classification problem.

Table 2 presents the extracted features for the four types of bio-signals of interest. Furthermore, in this module sensor behaviour is controlled and information feedback is provided for sensor operations. Fig. 2 presents a schematic representation of the module. The extracted features are analyzed below:

Mean value: It computes a vector of mean values over time for a specific input signal.

STD: It is the Standard deviation of a signal and results in a vector that could be described as the variation of the input signal around its mean value

Table 2: The Extracted features

ECG	EDA	EMG	Respiration
Mean Value	Mean Value	Mean Frequency	Mean Value
STD	STD	Median Frequency	STD
Mean abs fd	Mean abs fd	Mean Value	Mean abs fd
Mean abs sd	Mean abs sd	Root Mean Square	Mean abs sd
Mean amp	Mean Resp amp	STD	Mean amp
STD amp	STD Resp amp	Mean Value	STD amp
Mean Rise dur	Mean Rise dur		Rate
STD Rise dur	STD Rise dur		
Rate	Rate		

Mean and median frequency: They compute vectors of mean and median frequencies over time for a specific input signal.

Mean_abs_fd and Mean_abs_sd: For an acquired biosignal $X_N = (x_1, x_2, ..., x_N)$ the mean_abs_fd (mean absolute first difference) is defined as :

$$mean\_abs\_fd = \sum_{1}^{N} \frac{|x_2 - x_1| + |x_3 - x_2| + ... |x_N - x_{N-1}|}{N}, \qquad (1)$$

while the mean_abs_sd (mean absolute second difference) is defined as:

$$mean\_abs\_sd = \sum_{1}^{N} \frac{|x_3 - x_1| + |x_4 - x_2| + ... |x_N - x_{N-2}|}{N}, \qquad (2)$$

where x_i denotes a signal sample and N is the number of samples. These features are simple approximations of the first and second derivate respectively and therefore indicate fast changes in the recorded biosignals.

Mean amp and STD amplitude: They compute vectors of the mean amplitude and the Standard Deviation of the amplitude over time for the biosignals.

Mean Rise dur and STD rise dur: They compute vectors of the mean rise duration and the standard deviation over time.

Rate: It calculates vectors of the heart, respiration and EDR rate over time.

d) The facial representation module: The facial animation module models user's muscle contraction according to a 3-layer model, consisting of skull, muscle and skin layers. Each layer consists of a number of nodes, which are connected with neighbouring nodes of the same layer and nodes in the layers above/below. Each node represents a mass and each link between nodes is modelled as a spring.

The module flow goes through several processing stages before producing the 3D reconstruction:

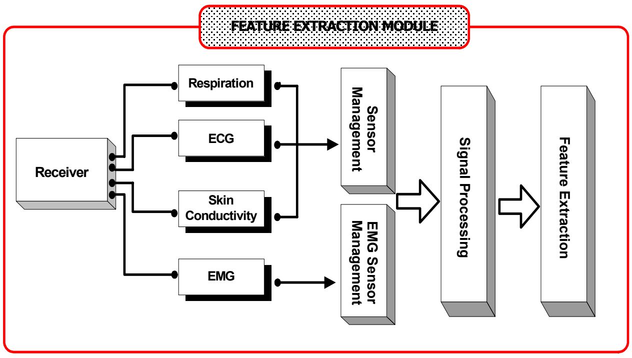

Figure 2: The Feature Extraction Module

(i) The features of the EMG signals, as extracted by the Feature Extraction Module, are used to estimate the contraction of the subject's monitored muscles. The outcome of this procedure is the quantification of muscle contraction for the 16 muscles being monitored, in the range 0.0 … 1.0 (where value 0 means that the muscle is not contracted and value 1 means that the muscle is fully contracted).

(ii) The contraction level drives the muscle model, to calculate the new position of muscle-nodes. The muscle model is simulating linear and sphincter muscles, which are the kinds of muscles examined in AUBADE.

(iii) Numerical methods, through the attachment of muscle nodes in the face's geometry, solve the mathematical model of the mass-spring network, given the new position of the muscle nodes.

(iv) The displacement of each node of the skin mesh is then applied to the face's geometry, as calculated by the mathematical model in the previous step. The resulting mesh is then presented on the user's screen as illustrated in Fig. 3.

e) AUBADE databases: The system's databases store the acquired raw signals which are ranked per user, date, event etc. They can be recalled any time from this database and can be analysed by specialists and researchers who are able to draw statistical and other information. The databases also store the medical history of the subjects, as well as their facial animation videos.

(f) The intelligent emotion recognition module: The Intelligent Emotion Recognition module is a decision support system that combines the extracted features in order to extract subject's basic emotions. A module's schematic is presented in Fig. 4. The classification into predefined emotional classes was achieved using Support Vector Machines (SVM) [24-25].

SVM is a new and very promising classification technique that can be seen as an alternative training

ISSN: 1727-1983 © 2005 IFMBE

technique for Polynomial, Radial Basis Function and Multi-Layer Perceptron classifiers. The main idea behind the technique is to separate the classes with a surface that maximizes the margin between them.

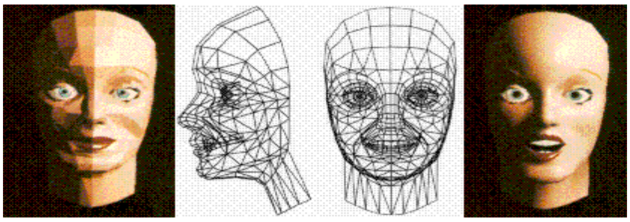

Figure 3: Resulting 3-D generic model presented to the user

A classification task based on SVM involves training and testing datasets which consist of a number of data instances. Each instance in the training set contains one "target value" (class labels) and several "attributes".

The goal of the SVM is to produce a model which predicts target value of data instances in the testing set in which only the attributes are given.

Consider a training set of instance-label pairs to be (x_i, y_i) , $i = 1,...,l$, where $x_i \in \Re^n$ is the training vector belonging to one of the emotional classes, l is the number of the extracted features in the training set and y_i indicates the class of x_i .

The support vector machine requires the solution of the following optimization problem:

$$\min_{w,b,\xi}\left(\frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{l}\xi_i\right), \tag{3}$$

subject to $y_i(\mathbf{w}^T\phi(\mathrm{x}_i) + b) \geq 1 - \xi_i$, $\xi_i \geq 0$, where *b* is the bias term, $\mathbf{w}$ is a vector perpendicular to the hyperplane < **w** ,b>, ξ is the factor of classification error and *C*>0 is the penalty parameter of the error term.

The training vectors x_i are mapped into a higher dimensional space *F* by the function $\phi : R^n \to F$.
SVM finds a separating hyperplane with the maximal geometric margin and minimal empirical risk R_{emp} in the higher dimensional space. R_{emp} is defined as:

$$R_{emp}(a) = \frac{1}{2l}\sum_{i=1}^{l}| y_i - f(x_i, a) |, \tag{4}$$

where *f* is the decision function defined as:

$$f(x) = \sum_{i=1}^{l} y_i a_i K(x_i, x) + b, \tag{5}$$

where $K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)$ is the kernel function, a_i are weighting factors and *b* is the bias term.
In our case the kernel is a radial basis function (RBF) which is defined as:

$$K(x_i, x_j) = \exp(-\gamma \| \mathrm{x}_i - x_j \|^2), \gamma > 0, \tag{6}$$

where $\gamma = \frac{1}{2\sigma^2}$ (σ is the standard deviation) is a kernel parameter. The RBF kernel non-linearly maps samples into a higher dimensional space, so it can handle the case when the relation between class labels and attributes is nonlinear.

Results

The system has been validated using data obtained from four drivers in simulated race conditions. An experienced psychologist supervised the whole procedure and annotated each driver's emotional state every 10 s.

The emotional classes identified were high stress, low stress, disappointment, euphoria and neutral face. The extracted vector of features along with the expert's annotation for every period of 10 s constituted the dataset for the classifier.

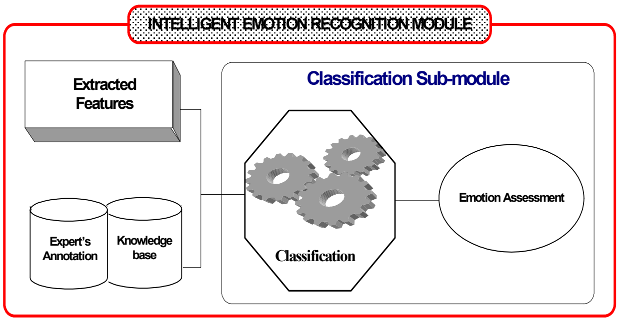

Figure 4: The Intelligent Emotion Recognition Module

The feature extraction and the classification of the emotional state have been exhaustively tested and validated for driver #1. The methodology followed was: a whole race was used for the training of the classifier and a different race was used for testing it.

The classifier is trained to identify emotion-specific patterns in a given dataset. Several classification methodologies have been applied in order to achieve an estimation of the subject's emotional state.

The classifier which provided the best results, as far as the accuracy is concerned, is the Support Vector Machine (SVM) classifier using RBF kernel.

The kernel's parameters γ and *C* were defined heuristically after a series of experiments and more specifically γ was equal to $2^{3.5}$ and *C* was equal to $2^{1.5}$. The averaged results are presented in Table 3 in terms of Sensitivity[1] and Positive Prediction Accuracy (PPA)[2].

1

$$\text{Sensitivity}_{\text{emotion\_a}} = \frac{\text{\# of templates classified as emotion\_a according to classifier}}{\text{total \# of templates belonging to class emotion\_a according to psychologist}} \times 100\%$$

2

$$\text{PPA\_emotion\_a} = \frac{\text{\# of templates correctly classified into class emotion\_a according to classifier}}{\text{sum of templates classified into class emotion\_a according to the classifier}} \times 100\%$$

ISSN: 1727-1983 © 2005 IFMBE

Table 3: Evaluation Results

%	**High stress**	**Low stress**	**Euphoria**
Sensitivity	91.3	84.4	82.6
PPA	86.7	73.6	87.9
	Disappointment	**Neutral face**	**Classification Rate**
Sensitivity	79.3	92.4	**86.0**
PPA	90.9	93.3	

Discussion

The AUBADE system recognizes and estimates basic emotions in real-time, in the form of a "diagnosis". AUBADE is a multifunctional system that can be utilized in many different ways and in multiple application fields.

The system's clinical application is based on the ability of supporting clinical diagnosis related to all the pathologies according to which the patient's capability to feel and express emotions is limited or totally absent. In those cases, doctors need to know the physiological condition of their patients. This is achieved by recording the expressions of the patient's face. Thus, muscle spasms as well as skin conductivity measurements are of key importance.

The goal is to evaluate the emotional state of the subject/patient. Possible patient categories include those in intensive care units of hospitals, those under suppression and patients that have hearing impairments. Moreover, after the administration of specific drugs, it is possible temporarily to normalize or even decrease the facial muscular activity.

Using the proposed system, we are able to follow the response of the patient to the specific drugs, adjusting and optimizing the dosages prescribed. Finally, applied to the intensive care unit (ICU) of a hospital, AUBADE will decrease the doctor's response time; the system undergoes the task of continuous tense monitoring of the patient. This will increase efficiency and save time.

As far as the car racing domain, AUBADE will be a useful tool for the mechanics of car racings, because they will be able to monitor emotionally the users. Moreover the car's setting and development will not only be based in subjective questionnaires filled by the driver, but in driver's emotional state (fear, stress level), which straightly correlates with the car's performance.

Finally, it may reduce accidents in car racings. Emotions and our psychological situation generally affect our behavior and reactions. Thus, if some emotion is detected that in some way may affect the behavior of the user, then the observer will be able to provide him with additional advices and guidance, preventing some reaction of the user that would be fateful.

AUBADE's classification accuracy into five predefined emotional classes is 86.0%. It must be noticed that the above results, although promising, are only indicative.

The system will be extensively tested and evaluated on car racing drivers of Maserati, following all relevant Federal Insurance Administration (FIA) regulations and other European ethical directives in relation to privacy of personal data and secure transfer of medical information.

Conclusions

A novel system that automatically monitors and classifies the psychological condition of human subjects from a set of emotions by applying pattern recognition techniques presented. AUBADE estimates the emotional state of human subjects by classifying vectors of features extracted from: facial Electromyogram, Respiration, Electrodermal Activity and Electrocardiogram. It is designed to be applicable to persons operating under extreme stress conditions, such as car racing drivers. In the medical field, AUBADE may be effectively utilized for patients suffering from neurological and psychological disorders.

The usual way to assess human emotion is by employing advance image-processing techniques in order to extract the facial characteristics. In our case, it is very difficult to apply image-processing techniques, since for safety reasons the users are wearing a mask and above it a casque. The proposed system realises an alternative method in order to record the facial expressions of the subject. Instead of using image-processing techniques, AUBADE utilizes the processing of surface EMG sensors, placed on the fireproof mask that the users are currently wearing.

A computational method for emotion recognition based on an SVM classifier is introduced. The method appears to have high performance both in accuracy and computational efficiency. Due to the fact that emotions vary from person to person, the system must be trained using a variety of drivers and then tested for its performance.

Acknowledgements

The work is part funded by the European Commission; project AUBADE (IST - 2002 - 507605).

References

[1] PICARD R.W. (2000): 'Affective Computing', MIT Press.

[2] RICHINS MARSHA L. (1997): 'Measuring Emotions in the Consumption Experience', Jpurnal of Consumer research, 24, pp. 127-146

[3] JENKINS J. M. , OATLEY K. and STEIN N. L. (1998): 'Human emotions: A Reader', Black-well Publishers, MA

[4] PICARD R. W., VYZAS E. and HEALEY J. (2001): 'Toward Machine Emotional Intelligence: Analysis of Affective Physiological State', IEEE Transactions Pattern Analysis and Machine Intelligence, 23, pp. 1175-1191.

[5] EKMAN P., FRIESEN W. (1975): 'Unmasking the Face'. Prentice- Hall

ISSN: 1727-1983

[6] BLACK M and YACCOB Y. (1997): 'Recognizing facial expressions in image sequences using local parameterized models of image motion', International Journal on Computer Vision, 25, pp.23-48
[7] ESSA I. and PENTLAND A.(1994): 'A vision system for observing and extracting facial action parameters', Proceedings of the CVPR'94, Seattle, Washington , USA.
[8] ESSA I. and GARDNER (1997): 'A. Prosody analysis for speaker affect determination' Proceedings of the workshop perceptual user interfaces, pp.45-46,
[9] BARLET M., HAGER J.C, EKMAN P. and SEJNOWSKI T.J (1999): 'Measuring Facial Expressions by Computer Image Analysis' Psychophysiology, 36, pp. 253-263
[10] DONATO G., BARLET M.S., Hager J.C., EKMAN P. and SEJNOWSKI T.J. (1999): 'Classifying Facial Actions', IEEE Trans. Pattern Analysis and Machine Intelligence,21, pp.974-989
[11] DESILVA L.C. , MIYSATO T. and NAKATSU R. (1997): 'Facial Emotion Recognition Using Multi-Modal Information', Proceedings of the IEEE Intelligent Conf. Information, Comm. And Signal Processing, pp. 397-401
[12] HUANG T.S. CHEN L.S. and TAO H., (1998): 'Bimodal Emotion Recognition by man and Machine' Proceedings of the ATR workshop virtual communication Enviroments
[13] CHEN L.S., HUANG T.S., MIYSATO T. And NAKATSU R., (1998):'Multimodal Human emotion/Expression Recognition' Proceedings of the 3rd International Conference Automatic Face and Gesture Recognition
[14] CHEN. L.S., (2000): 'Joint processing of audio-visual information for the recognition of emotional expressions in human-computer interaction', PhD thesis, University of Illinois at Urbana-Champaign, Dept. of Electrical Engineering
[15] COHEN I. , SEBE N. , COZMAN F. , CIRELO M., and HUANG T.S. , (2003): 'Learning bayesian network classifiers for facial expression recognition using both labeled and unlabeled data', In Proc. Conf. on Computer Vision and Pattern Recognition, 1, pp. 595–601
[16] COHEN I., SEBE N, GARG A. , CHEN L., and HUANG T.S. (2003): 'Facial expression recognition from video sequences: Temporal and static modeling', Computer Vision and Image Understanding, 91, pp. 160–187
[17] OLIVER N., PENTLAND A., and BERARD F. , (2000): 'A real-time face and lips tracker with facial expression recognition', Pattern Recognition, 33, pp. 1369–1382
[18] CACIOPPO J.T. and TASSINARY L.G., (1990): 'Inferring Physiological Significance from Physiological Signals', American Psychologist, pp.16-28.
[19] EKMAN P., LEVENSON R.W. and FRIESEN W.V.,(1983): 'Autonomic Nervous system Activity Distinguishes Among Emotions', Science, 221, pp.1208-1210.
[20] WINTON W.M. PUTNAM L. and KRAUSS R., (1984): 'Facial and Autonomic Manifestations of the Dimensional Stucture of Emotion', Journal Experimental Social Psychology, 20, pp.195-216.
[21] FLIDLUND A.J. and IZARD E.Z., (1983): 'Electromyographic studies of facial expressions of emotions and patterns of emotions'. Social Phychophysiology: A sourcebook Cacioppo and R.E. Petty eds.
[22] CACIOPPO J.T. , BERNTSON J.T., LARSEN J.T. POEHLMANN K.M. and ITO T.A.,(2000): 'The Psychophysiology of emotion'. Handbook of emotions. Lewis M. and Haviland-Jones eds.
[23] HEALEY J.A.,(2000): 'Wearable and Automotive Systems for affect recognition from Physiology', PhD Thesis, Massachusetts Institute of Technology,
[24] BOSER B.,GYUON I. and VAPNIK V. (1992) 'A training algorithm for optimal margin classifiers', On Proceedings of the 5th Annual Workshop on Computational Learning Theory.
[25] CORTES C. and VAPNIK V. , (1995): 'Support-vector network' Machine Learning,32, pp.273-297

ISSN: 1727-1983