# Biometric Valence and Arousal Recognition

Christian Martyn Jones
University of the Sunshine Coast
Queensland 4558
Australia
+61 7 5459 4849
cmjones@usc.edu.au

Tommy Troen
Heriot-Watt University
Edinburgh EH14 4AS
United Kingdom
+44 131 449 5111
tt16@macs.hw.ac.uk

## ABSTRACT
A real-time user-independent emotion detection system using physiological signals has been developed. The system has the ability to classify affective states into 2-dimensions using valence and arousal. Each dimension ranges from 1 to 5 giving a total of 25 possible affective regions. Physiological signals were measured using 3 biometric sensors for Blood Volume Pulse (BVP), Skin Conductance (SC) and Respiration (RESP). Two emotion inducing experiments were conducted to acquire physiological data from 13 subjects. The data from 10 of these subjects were used to train the system, while the remaining 3 datasets were used to test the performance of the system. A recognition rate of 62% for valence and 67% for arousal was achieved within +/- 1 units of the valence and arousal rating.

### Categories and Subject Descriptors
B.4.2 [**Input/Output and Data Communications**]: Input/Output Devices - Channels and controllers. C.3 [**Computer System Organisation**]: Special-Purpose and Application-based Systems - Real-time and embedded systems, Signal processing systems. H.1.2 [**Models and Principles**]: User/Machine Systems - Human factors, Human information processing. H.5 [**Information Systems**]: Information Interfaces and Presentation - Input devices and strategies

### General Terms
Algorithms, Measurement, Experimentation, Human Factors

### Keywords
Affective Computing, Emotion Recognition, Biometrics, Valence and Arousal

## 1. INTRODUCTION
Emotions are one of many unique and complex human attributes that influence people's day to day behavior and their interaction with others. Research in neuroscience, psychology and cognitive science suggests that emotion plays a critical role in rational and intelligent behavior, and that emotional skills are especially important for learning preferences and adapting to what is important [1-3].

Emotional expressions are conveyed through a variety of different channels, often concurrently. People express their emotions through facial expressions, body movement, gestures and tone of voice, and expect others to detect, understand and respond to their affective state. Much of the research on emotion detection in the field of computer science has traditionally been done with assessment of physical factors such as speech, facial expressions and a combination of the two (bimodal) [4].



OzCHI 2007 Proceedings, ISBN 978-1-59593-872-5

However there is a distinction between inner emotional experiences and the outward emotional expressions that people use to convey messages about their emotional states [5]. Some emotions can be hard to recognise by humans, and inner emotional experiences may not be expressed outwardly. The question remains whether physiological patterns can be used to recognise distinct emotions [6]. In recent years methods for emotion detection using physiological signals have been extensively investigated [1], providing encouraging results where affective states are directly related to change in inner bodily signals

Affect is a broad definition that includes feelings, moods, sentiments etc. and is commonly used to define the concept of emotion [5]. There are two different perspectives to consider when labeling an affective state i) the discrete emotion model where predefined labels are used to describe an emotional experience (e.g. happy, sad, etc) or; ii) the dimensional model where the affective state is determined by arousal and valence. Valence can be described as a subjective feeling of pleasantness or unpleasantness while arousal is the subjective state feeling activated or deactivated [7].

Affective states don't easily map into distinctive emotional labels, however many researchers agreement to at least six universal categories of happiness, sadness, fear, anger, surprise and disgust [8]. Although there are different opinions on how to label emotions, most researchers believe that valence and arousal play a role in the measurement of affective states and that one dimension rarely is sufficient to fully describe affect [7][9]. Many models have been developed to describe the two-dimensional space of arousal and valence, such as the Circumplex model by Russel [10] where the two dimensions are represented by a vertical arousal axis (ranging from deactivation to activation) and a horizontal valence (ranging from unpleasant to pleasant) axis. These models can provide a mapping between predefined labels and the level of arousal and valence [8], Figure 1.

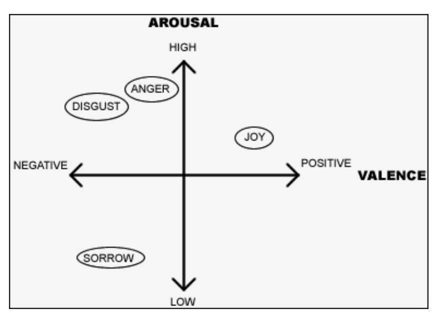


**Figure 1. Adapted Circumplex model of affect showing location of emotional labels.**

## 2. RECOGNISING EMOTION FROM PHYSIOLOGICAL SIGNALS
The performance of emotion recognition systems using physiological signals can vary depending of the range of number of emotional categories and whether the systems are

user dependent or independent. Picard et al. have developed a physiological emotion recognition system able to recognise eight emotions of anger, hate, grief, platonic love, romantic love, neutral state, joy and reverence with 81% accuracy for a single user [1]. Physiological signals from 4 biometric sensors were collected over a period of 20 days. Statistical features were then calculated over one-day periods, selected and classified with a hybrid of Sequential Floating Forward Search (SFFS) and Fisher Projection (FP).

Kim et al. developed a user-independent emotion detection system using 3 physiological signals [11]. The physiological data were collected from 50 children aged 5 to 8 years old to recognise four emotional states of sadness, anger, stress and surprise. A support vector machine (SVM) was utilized as a pattern classifier. Recognition rates of 78.43% and 61.76% were achieved for three and four emotional states respectively. Adults and children tend to react to emotion differently so one might conclude that this system is user-independent within the population (children) it is trained on.

Nasoz et. al have conducted a study to model user emotional state [12]. Three physiological measurements were used and data collected from 31 participants (male and female) from a student population. This study used normalized signals instead of statistical features and employed two separate classification methods: k-nearest-neighbour (KNN) and Discriminant Function Analysis (DFA). The best results were achieved using DFA to classify 5 emotions with 90% accuracy for fear, 87.5% for sadness, 78.58% for anger, 56.25% for surprise and 50% for frustration.

Healey et al. have used 4 physiological signals to detect the intensity of stress in automobile drivers [13]. Sequential Forward Floating Search (SFFS) was used to recognize patterns of stress whereby the intensity of stress was recognized with 88.6% accuracy. Fernandez used two physiological signals to detect frustration in computer users [14]. Statistical features and Hidden Markov models were used to detect the level of frustration in users that were interacting with poorly implemented computer programs.

Our work concentrates on developing a user independent emotion recognition system which can classify affective state in 2-dimensional valence/arousal space representing 25 affective regions.

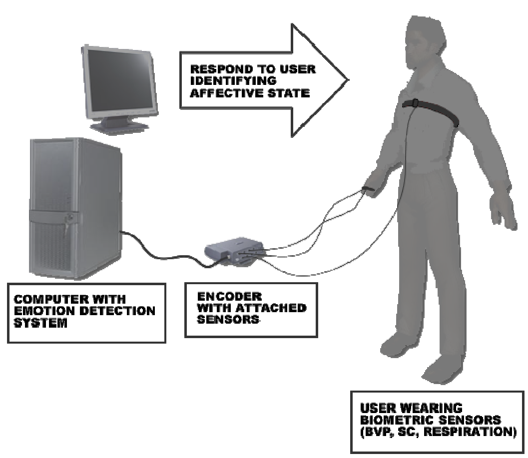


**Figure 2. System implementation**

# 3. SYSTEM IMPLEMENTATION

This project has developed a user independent biometric emotion detection system which can automatically detect and recognise user's emotional state in real-time. The system uses Blood Volume Pulse (BVP), Skin Conductance (SC) and Respiration (RESP) sensors, to classify human affective states in 2-dimensional valence/arousal space, Figure 2.

The system has four modules, Figure 3:

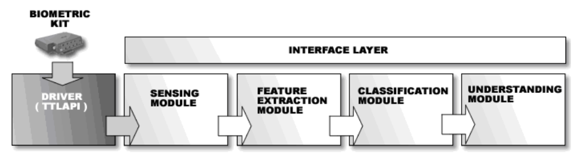


**Figure 3. System modules for emotion recognition**

1. *sensing module* to read in physiological signals from the 3 biometric sensors and process them. The Blood Volume Pulse (BVP) signal is a relative measure of the amount of blood flowing in a vessel. From BVP we calculated heart rate and heart rate variability. The heart rate is known to reflect emotional activity and has been used to differentiate between both negative and positive emotions as well as different arousal levels [15]. We use the Skin Conductance (SC) sensor to measure the skin's conductance (between two electrodes and is a function of sweat gland activity and the skin's pore size). As a person becomes more or less stressed, the skin's conductance increases or decreases proportionally [5]. We use the respiration (RESP) signal (a relative measure of chest expansion) to calculate the respiration rate and relative breath amplitude [5]. Emotional arousal increases respiration rate while rest and relaxation decreases respiration rate, negative emotions generally cause irregularities in the respiration pattern [16]. The sensor module includes a Thought Technology ProComp Infinity encoder [15] connected to PC with a USB cable.
2. *feature extraction module* to extract features from the raw data to provide a closer correlation between physiological changes and emotions than using the raw data itself. Residing on the PC, the module calculates 11 features (e.g. standard deviation and mean of the raw signals, heart-rate and heart-rate acceleration etc.) before passing these features on to the classification module.
3. *classification model* to recognise and classify an affective state on the basis of features calculated from the 3 physiological signals. The classifier resides on the PC and consist of two artificial neural networks (one Multi Layer Perceptron (MLP) for each affective dimension), which recognise patterns from inputs and classify these patterns into 2-dimensional affective states.
4. *understanding module* to presented a graphical presentation of the emotion recognition to the user. The system is designed and implemented to allow easy integration with applications such as games, and online learning applications etc. where emotion recognition can offer adaptive control to maintain user interest and engagement. Once connected via sensors to the emotion recognition system, the affective state of the user is displayed continuously and in real-time on a displayed 2-dimenstional graph of valance and arousal, Figure 1, and as highlighted characters representing valence and arousal, Figure 5.

## 4. ACQUISITION OF TRAINING AND TEST DATA

Physiological data from users in a range of valence/arousal affective states is required to train and test the biometric emotion recognition system. 13 participants (3 male, 10 female, aged 19-35) were wired to the sensor module which recorded 256 samples of BVP, SC and RESP per second. The sensors were attached to the participant's left hand (the right hand free for participant to rate pictures), Figure 4. They were asked to keep their left hand (with sensors attached) as still as possible to prevent any signal noise from the sensors. Each participant was left alone in the laboratory during the session which lasted for about 15 min.

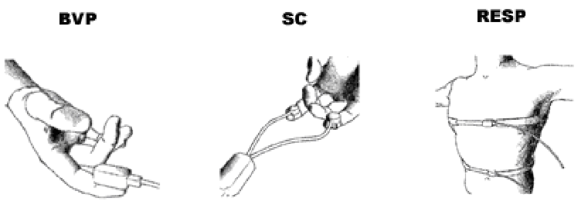


**Figure 4. Sensor placements [16].**

Participants viewed a slideshow of 21 images assembled from the International Affective Picture System (IAPS) [17]. IAPS consist of 6000 pictures designed to induce emotions in people and each picture has been pre-rated for valence and arousal. The selected 21 images were set up in a pseudo-randomized order of a series of 3 for each dimension eg first an unpleasant rated picture, then a neutral rated picture and finally a pleasant rated picture. Each picture was viewed for 15 seconds, after which there was a rest period for 25 seconds where the participants were asked to conduct their own subjective rating of how the picture made them feel.

The participant rates their emotion using an in-house application, Figure 5, based on the Self-Assessment Manikin (SAM) affective rating system devised by P. J. Lang [17]. The rating system consisted of two 5-point scales: a valence scale ranging from unhappy to happy and an arousal-scale ranging from calm to excited, giving a total of 25 possible combinations. All ratings were stored in a database to correlate how the participants felt when viewing the picture with the original ratings from IAPS. This information was also later used together with the IAPS ratings to train the emotion recognition system.

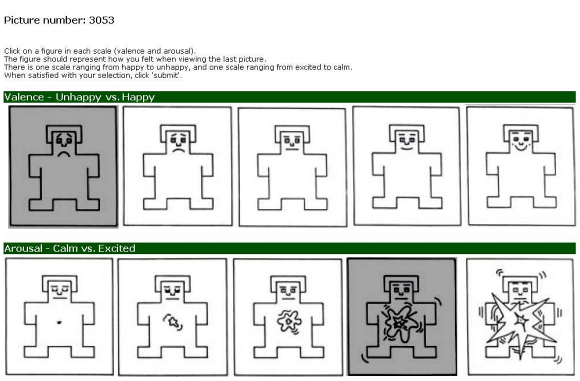


**Figure 5. Subjective picture rating application based on the Self-Assessment Manikin [17]. Participant has rated current picture as unhappy (position 1 on valence) and interested (position 5 on arousal)**

## 5. TRAINING EMOTION CLASSIFIERS

The two neural networks hereby referred to as the valence network and the arousal network, have been trained on 10 of the13 collected datasets. Each dataset was divided into sample segments for each emotion inducing picture in the experiments. Features are extracted from a segment of samples, where each segment is 5 seconds long giving a total of 1280 samples per segment. Sample data from each sensor is retrieved from the sensor module every 5 seconds and the features are then calculated and form the feature vector containing 11 values. Each viewed picture has an associated valence and arousal rating from IAPS and a subjective rating given by the participants. An average over all participants was created for the subjective ratings for each picture, and these ratings were again averaged with the IAPS ratings forming the final ratings used for the training and testing of the system, Table 1.

**Table 1. Proportion of 21 emotive pictures corresponding to each affective state in the valence and arousal dimensions (1 to 5 represents unhappy to happy in the valence dimension and calm to excited in the arousal dimension)**

| Number of pictures | | |
|---|---|---|
| **Affective State** | **Valence** | **Arousal** |
| 1 | 5 | 4 |
| 2 | 4 | 6 |
| 3 | 5 | 4 |
| 4 | 5 | 3 |
| 5 | 2 | 4 |
| *Total* | *21* | *21* |

Having extracted and standardized features for all 10 datasets for 21 segments the total number of training patterns should be 210. However the sample data from 6 of the segments for one of the datasets was corrupt and had to be disregarded, therefore allowing 204 training patterns. Each of the patterns corresponds to a valence and arousal rating which is referred to as the outputs. A training file was generated for the valence network where each pattern was accompanied by its corresponding valence output. Another training file was generated for the arousal network using the same patterns however these were accompanied by their corresponding arousal outputs.

## 6. RECOGNITION PERFORMANCE

Of the 13 datasets containing physiological signals collected, 10 were used to train the classifiers, providing 3 datasets to test the recognition performance. Each dataset was divided into 21 sample segments corresponding to the final ratings of the emotion inducing pictures, forming a testing set consisting of 63 sample segments. The testing set was then processed by the emotion classifiers in an offline mode undergoing all the same operations as in a real-time run. The classifiers returned outputs valence and arousal ratings for each segment.

Two criterions were used in the performance test: (1) the number of times the system produced outputs which were an exact match to the final ratings for each segment and; (2) the number of times the outputs produced were within +/- 1 unit of the final ratings for each segment (as people find it difficult to separate between two very closely related emotional states, such the intensity of 2 and 3 in the arousal or valence dimension). For criteria (1) an overall recognition rate of 30%

and 35% were achieved for the valence and arousal dimensions respectively, Table 2 (where guess rate is 60% if correct value is 2, 3 or 4, and 40% if correct value is 1 or 5). For criteria (2) with a deviation of 1 unit in each dimension (i.e. the system output is very close to the true affective state but not an exact match) recognition rates of 62% for valence and 67% for arousal were achieved, Table 2.

**Table 2. Valence and arousal classification recognition rates with test dataset**

| RECOGNITION RATES | | |
|---|---|---|
| | **Exact match in %** | **With deviation +/- 1 in %** |
| **Valence** | 30 | 62 |
| **Arousal** | 35 | 67 |

The recognition rate for each affective state in the valence and arousal dimensions as shown in Table 3.

**Table 3. Recognition rates for each affective state with the test dataset**

| | **VALENCE RECOGNITION RATE** | | **AROUSAL RECOGNITION RATE** | |
|---|---|---|---|---|
| **Rating** | **Correct in %** | **With +/- 1 in %** | **Correct in %** | **With +/- 1 in %** |
| 1 | 53 | 67 | 17 | 75 |
| 2 | 8 | 75 | 67 | 89 |
| 3 | 40 | 67 | 17 | 83 |
| 4 | 27 | 60 | 22 | 22 |
| 5 | 0 | 17 | 33 | 42 |

Low rated valence states such as unhappy can be classified with 53% accuracy and neutral states with 40% accuracy, confirming that the system is able to discriminate negative valence emotions and neutral states. However discrimination of positive valence states appears more difficult. Of the 21 pictures used in the experiments to train the system, only 2 had a valence rating of 5. Positive valence states are more difficult to elicit in subjects using pictures than negative valence states. An investigation of the changes in the physiological signals of a subject when experiencing negative and positive states showed that the changes in the physiological signals where more significant when experiencing negative than positive valence states. A higher recognition rate for positive valence states could be achieved if more suitable stimuli for positive affective states is used.

The recognition rates for the various states in the arousal dimension shows that the system is able to recognise low arousal states with an accuracy as high as 67%, while higher arousal states are recognised with an accuracy of 33%. One may assume that high arousal would be more easily detected with a higher accuracy than low arousal. It is known that physiological signals such as the heart-rate will vary greatly when a person is experiencing high arousal states such as stress. One would naturally believe that these changes would provide a good discrimination between the high aroused state and the other affective states, making it easier for the classifier to recognise a pattern. However the pictures may not be sufficient to elicit high arousal in participants and therefore not inducing physiological changes that would usually occur when a person is excited or stressed.

# 7. CONCLUSIONS

This paper presents a fully implemented user-independent biometric emotion detection system. The system can successfully recognize 25 affective states in 2-dimensional space using physiological data from 3 sensors. A flexible real-time emotion detection framework have been developed allowing the system to be integrated into other application offer emotionally responsive and intelligent systems. The overall system performance within +/- 1 of user and IAPS ratings is 62% for valence and 67% for arousal.

# 8. REFERENCES


**[1]** Vyzas, E., Picard, R.W. and Healey, J. Toward Machine Emotional Intelligence: Analysis of Affective Physiological State, *IEEE Transactions on Pattern Analysis and Machine Intelligence, 23, 10* (Oct 2001)

**[2]** Salovey, P. and Mayer, J.D. Emotional Intelligence, *Imagination, Cognition and Personality, 9, 3* (1990) 185-211

**[3]** Goleman, D., *Emotional Intelligence*, New York : Bantam Books, 1995

**[4]** Cowie, R., Douglas-Cowie, E., Tsapsoulis, N. and Votsis. G. Emotion recognition in human computer interaction. *IEEE Signal Processing Magazine*, (Jan 2001), 32-80

**[5]** Picard, R.W. *Affective Computing*. Cambridge, MA: MIT Press, 1997

**[6]** Cacioppo, J.T. and Tassinary, L.G. Inferring Psychological Significance from Physiological Signals, *Am. Psychologist*, *45* (Jan 1990) 16-28

**[7]** Feldman B.L. *Discrete Emotions or Dimensions? The Role of Valence Focus and Arousal Focus.* (1998)

**[8]** Zagalo, N., Barker and A. Branco, V. *Story Reaction Structures to Emotion Detection* (2004)

**[9]** Feldman B.L. and Russel, J.A. *The Structure of Current Affect: Controversies and Emerging Consensus* (1999)

**[10]** Russel, J.A. *A Circumplex Model of Affect* (1980)

**[11]** Kim, K.H., Bang, S.W. and Kim, S.R.. *Emotion recognition system using short-term monitoring of physiological signals* (2004)

**[12]** Nasoz, F., Lisetti, C.L., Alvarez, K. and Finkelstein, N. Emotion recognition from Physiological Signals for User Modelling of Affect *Proc of the 3rd Workshop on Affective and Attitude user Modelling*, Pittsburgh, PA, 2003

**[13]** Healey, J. and Picard, R. W. SmartCar: Detecting Driver Stress *Proceedings of ICPR'00*, Barcelona, Spain, 2000

**[14]** Fernandez, R. *Stochastic Modelling of Physiological Signals with Hidden Markov Models: A Step Toward Frustration Detection in Human-Computer Interfaces*, M.Sc. Thesis, MIT, 1997

**[15]** Thought Technology Ltd. "*ProComp Infinity Hardware* Manual", www.thoughttechnology.com/manual.htm

**[16]** Stern, R. M., Ray, W. J. and Quigley, K.S., *Psychophysiological Recording*. New York: Oxford University Press 2001

**[17]** Lang, P. J., Bradley, M. M. and Cuthbert, B. N.. *International affective picture system (IAPS): Technical Manual and Affective Ratings*, Center for Research in Psychophysiology, University of Florida 1999