http://dx.doi.org/10.14236/ewic/HCI2018.221

Intelligent Joystick Sensing the User's Emotion and Providing Biofeedback

Roya Haratian
Bournemouth University
Poole, UK, BH12 5BB
rharatian@bournemouth.ac.uk

Development of an intelligent joystick is proposed which senses the user's bio-signals and recognises the user's emotion. It provides biofeedback to the user as well as the user's emotional state information to the computer allowing human-computer interaction over sensitive environment. While the user is interacting with a computer via a joystick the bio-signals can be collected through the user's fingers touching it. The collected bio-signals information is mapped on a two-dimensional space to find out the quality and intensity of emotion continuously and in a real-time manner. The intelligent joystick has application within several fields such as healthcare, sport and game industries. In such cases, the user can be influenced, or suffer from medical problems while under stress during interaction with the machines. The intelligent joystick will provide feedback to the user and alert alarm about unhealthy conditions through the embedded actuators and allow the machine to adapt with the users' emotional state.

Intelligent Systems, Human Computer Interaction, Bio-Signals, Biofeedback

1. INTRODUCTION

Sensing user's bio-signals for emotion recognition and providing biofeedback to the user through a joystick allow human-computer interaction over sensitive and responsive environment to human needs and emotions. While the user is interacting with a computer through a joystick as a common input device using in aircraft, game, and healthcare industries to control the system, the bio-signals can be collected through the user fingers connecting to the joystick as suggested in *Figure1*. The bio-signals information will be mapped on a two-dimensional emotion space to find out the quality and intensity of emotion continuously and in real-time manner using advanced Machine Learning approaches. Biofeedback will be provided to the user which can be in forms of sound, light, vibration, image or video.

Detection and recognition of emotions as affective states plays an important role in physical and mental health. As affective state and stress are related to each other, detection of emotions is used for accurately detecting stress. Stress has effect on one's physical conditions such as emotional distress, overall arousal, muscular ache and tension which can lead to heart attacks and possible sudden death (Greene et al. 2016). The corresponding physiological changes in blood volume pulse, skin conductance, and heart rate can be detected through biosensors (Acampora et al. 2013). By using advanced algorithms to detect the related emotion and level of stress in real-time manner, biofeedback can be provided to the user to alert about unhealthy conditions.

The affect can be described through three major approaches: categorical, dimensional, and appraisal-based. In categorical approach, emotions are described in six basic emotions such as happiness, sadness, surprise, fear, anger, and disgust. In a dimensional approach, emotions are described through two dimensions: valence and arousal. The valence refers how positive or negative the emotion is and ranges from unpleasant feeling to pleasant feeling. The arousal dimension refers to level of control, and ranges from low control to high control. In the appraisal model quality, intensity and duration of emotion are considered as well (Klaus 2005). The combination of dimensional and categorical models (Scherer 2013) with ratings of intensity for a number of major emotion families, Genova Emotion Wheel (GEW), is used in this paper to model emotion which assists continuous emotion modelling and detection in real-time manner so the duration is also considered.

To identify emotions, physiological or bio-signals are used in affect sensing collected through biosensors. Galvanic Skin Response, GSR, provides measurement of skin conductance which increases linearly with a person's level of overall

Proceedings of British HCI 2018. Belfast, UK.

arousal or stress. The GSR sensor measures skin conductance by applying a small voltage to the skin across the electrodes and the resulting current is measured. Blood Volume Pulse, BVP, is indicator of blood flow which can be measured by pulse sensor. Heart rate increases with negatively valence emotions such as anxiety or fear. A low heart rate can indicate a state of relaxation, whereas an increased heart rate can indicate a potential state of mental stress or frustration.

Respiration rate changes based on how deep and fast a person is breathing. Slow and deep breathing indicates a relaxed resting state while irregular rhythm, quick variations, and cessation of respiration corresponds to more aroused emotions like anger or fear. SpO2 sensor measures amount of oxygen in blood by emitting and then absorbing a light wave passing through blood vessels in the fingertip. A variation of the light wave passing through the finger will give the value of the SpO2 measurement because the degree of oxygen saturation causes variations in the blood's colour (Gunes et al, 2010).

Through the correlation between bio-signals and affective states the user's emotion can be recognized using Machine Learning algorithms. As the intensity of arousal is linearly correlated with skin conductivity and valence is correlated negatively with blood volume pulse rate and the muscle activity, we can extract arousal relevant features from GSR and valance relevant features from BVP, and respiration related signals such as SpO2.

In this paper after reviewing the literature in Section 2, the suggested methodology is discussed in section 3 including the proposed signal processing approach as well as model training procedure. The paper will be concluded in Section 4.

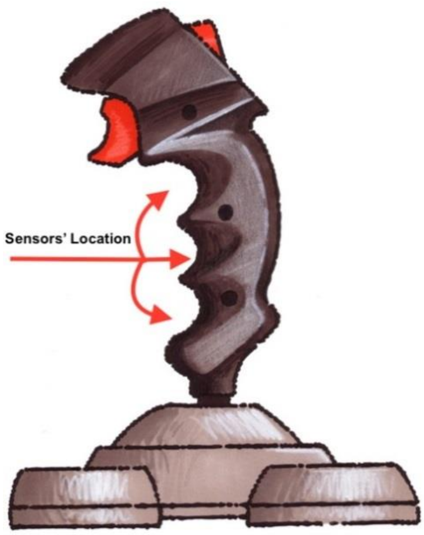

Figure 1*: The proposed location of sensors in a joystick, modified from (Kreativebomb.com, 2011)*

2. LITERATURE REVIEW

Different modalities have been used in literature for affect detection such as facial expressions, voice, body language and posture, physiology, brain imaging and text as well as integration of information from different modalities (Calvo et al, 2010). As in this paper physiological signals are the selected modalities, the literature review is focused on these modalities.

The literature mostly concerns with static emotion recognition considering discrete emotion states. In (Picard et al. 2001), electromyogram, blood volume pressure, skin conductivity and respiration rate are measured and by considering certain window of time series of the bio signals 8 different discrete emotions are predicted. In a more recent paper (Wanhui et al. 2014), by using advanced classification techniques emotions such as amusement, anger, grief, and fear are detected via fingertip physiological signals.

Detecting emotion in the two-dimensional emotion space with valence-arousal axes rather than discrete emotion states such as the reviewed papers is presented in the literature as well. The approach is either using classification techniques or recognition of valance and arousal separately.

Dividing the arousal and valence axes of emotion space into some levels and mapping emotions in that discrete classes using physiological signals such as heart rate, respiration rate and facial muscle contraction is introduced in (Kulic and Croft 2007). In (Kim and André 2008) similar physiological signals are used and the emotions are categorised into 4 different classes by dividing valence and arousal axes into 2 levels. A more recent research (Valenza et al. 2012) using electrocardiogram, electro-dermal response and respiration signals also classifies emotions into 16 classes by considering 4 arousal and valence discrete categories. By using discriminant classifier in (Nardelli et al. 2015) heart rate signals are mapped into four arousal and two valence classes.

Detecting stressful situation through the recognized emotions by mapping physiological signals such as skin temperature, cardiovascular, electro-dermal, breathing and facial muscle activity is introduced in (Santos et al. 2016). A continuous emotion recognition algorithm is introduced in (Nicolaou et al, 2011) using fusion of facial expression, gesture and audio cues. In the paper, level of arousal and valance is predicted over time continuously. Using long short-term memory (LSTM) Neural Network as building unit for layers of a recurrent neural network (RNN), the longer temporal dependency features are learnt in the algorithm. It enables continues prediction of valance and arousal over time. In (Jenke et al, 2018), a dynamic model is proposed based on dynamic field theory (DFT) enabling prediction of emotion intensity. DFT can be seen as a generalization of recurrent neural networks to continuous dimensions, adding a functional interpretation to each layer.

The contribution in this paper to the research field in comparison to the reviewed literature focusing on emotion recognition through physiological signals are in: modality, use of the most recent introduced emotion model in the literature, real-time and continues emotion recognition through the introduced algorithm and hardware development as list bellow with more details.

- Modality: physiological signals such as galvanic skin response, heart rate, respiration related signal (SpO2) that can be collected through fingers touching joystick rather than wide range of physiological signals
- Emotion space: the combination of dimensional and categorical models where the quality and intensity of emotion are present on a two dimensional space (Scherer 2013) as shown in *Figure 2*.
- Algorithm: recognising emotion in a real-time and continues manner by using advanced Machine Learning such as Dynamic Neural Fields approach.
- Hardware: development of a joystick where the embedded electronic system can sense the physio-signals while the user is touching it and providing visual, haptic, and auditory biofeedback .

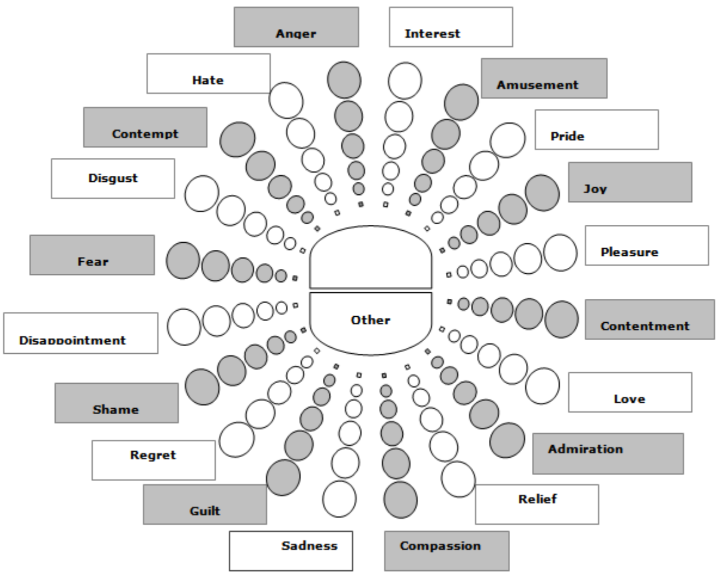

Figure 2: *The emotion model (Scherer, 2013)*

3. METHODOLOGY

The embedded biosensors in the joystick collects galvanic skin response, SpO2 and pulse rate. The signals are processed in a processing unit based on the emotion recognition algorithm which is described in this section. Biofeedback is provided to the user based on the sensed emotion in forms of visual (light or image), haptic (vibration) and auditory (alarm and sound) as the block diagram is shown in *Figure 3*.

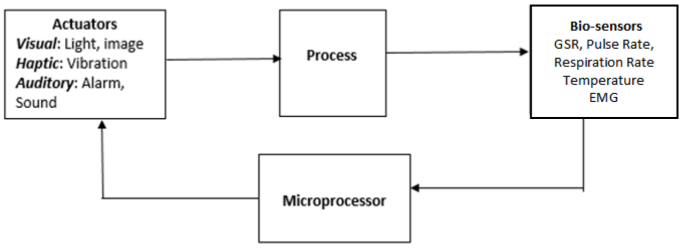

Figure 3: *The block diagram of the system*

For the purpose of emotion recognition, machine learning approaches (Bishop 2006) are suggested in the literature: Regression approaches such as Support Vector Machine or Deep Learning such as Neural Networks, Recursive neural networks, and long short-term memory (Russell and Norvig 2003). In this paper, the proposed model for recognition of emotion is based on Dynamic Neural Filed (DNF) which is generalization of recurrent neural networks to continuous dimensions which adds a functional interpretation to each layer. The proposed neural population is shown in *Figure 4*.

As the intensity of arousal is linearly correlated with skin conductivity and valence is negatively correlated with pulse rate, the GSR and pulse rate are inputs of the network. In the proposed DNF, there are two one dimensional input neurons, U1 and U2, where the two bio-signals are the inputs to them. The model inputs are GSR which is correlated with arousal and Pulse rate which is correlated with valance in the emotion model. To stabilise the signals there are two memory neurons, M1 and M2, where their outputs are the inputs of the emotion neuron, E. The two-dimensional emotion neuron is modelling the emotion space with valance and arousal dimensions (Haratian, 2018).

For the purpose of model training, a series of experiments are designed. The designed experiment consists of several trials of showing series of pictures which are rated by subjects while physiological signals are collected. The pictures are from IAPS, International Affective Picture System, database designed to provide a standardized set of pictures for studying emotion and attention. The dataset has been widely used in psychological research and was developed by the National Institute of Mental Health Centre for Emotion and Attention at the University of Florida (Haratian, 2018).

During the time that the picture is on the screen, the rating is made on a 2D labelling system (Haratian, 2016). Each series of pictures belongs to one emotion quality labelled on the outer borders of the 2D labelling system. Subjects can rotate a knob to the emotion quality label and move the indicator on radial line of the circle according to the emotion intensity they feel during viewing the pictures, the closer to the centre, the less intense emotion felt during viewing the picture based on GEW emotion model. The time length of showing each picture is 15s and it takes about 10s for an emotion to be induced using pictures. By the end of each trail a black screen is shown to help the subject reach neutral state for the felt emotion.

For the purpose of the model training, a mechanical emotion wheel is developed to collect self-reports of users in two dimensional emotion space by rotating and pushing or pulling the knob of the wheel in the appropriate direction. Emotion slider to collect self-report of users in one dimension by pushing or pulling the knob of the slider to measure the intensity of emotion has been introduced in literature. The advantage of current emotion wheel on other emotion measurement instrument in literature is that the quality and intensity of emotion can be labelled simultaneously in real-time.

The labels are similar to the Geneva Emotion Wheel, 3rd version labels. The 2D labelling system designed for our experiment (*Figure 5*) comprises a rotatory potentiometer which measures the angle by rotation of the knob and a slide potentiometer which measures the intensity of emotion by pushing or pulling the knob.

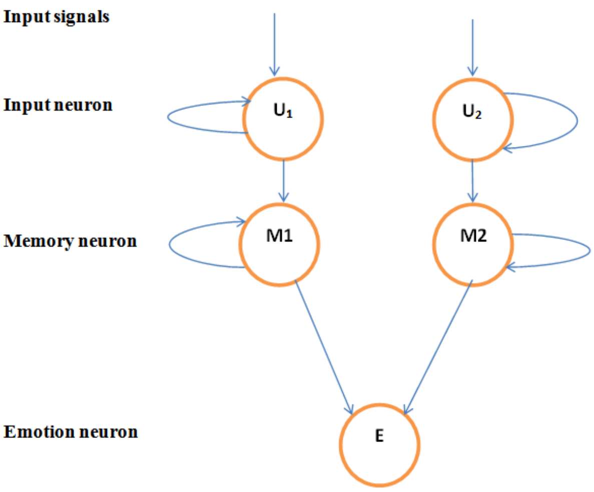

Figure 4: *The proposed neural population for the model*

The model parameters are determined through the training phase of the DNF, by minimizing a cost function. The DNF estimates the quality and intensity of emotion based on the input training signals (GSR and PR) in multiple trials. The cost function represents the estimation error based on the training data set where the actual quality and intensity of the emotion associated with the physiologic measurements are known. The cost function is calculated using equations (1) and (2) as the average of all mean square errors, MSEs, across multiple trials, for both quality and intensity of emotion (Haratian, 2018).

$$\boldsymbol{Cost} = \frac{1}{T*L}\sum_l\sum_t\left|\widehat{Q}_l(t) - Q_l(t)\right|^2 \qquad (1)$$

$$\boldsymbol{Cost} = \frac{1}{T*L}\sum_l\sum_t\left|\hat{I}_l(t) - I_l(t)\right|^2 \qquad (2)$$

In this equation, $\widehat{Q}(t)$ is the emotion quality estimate by the model and $Q(t)$ is quality label (target). $\hat{I}(t)$ is the emotion intensity estimate by model and $I(t)$ is intensity label (target). Total recording time for each trial is represented by *T* and number of trials by *L*.

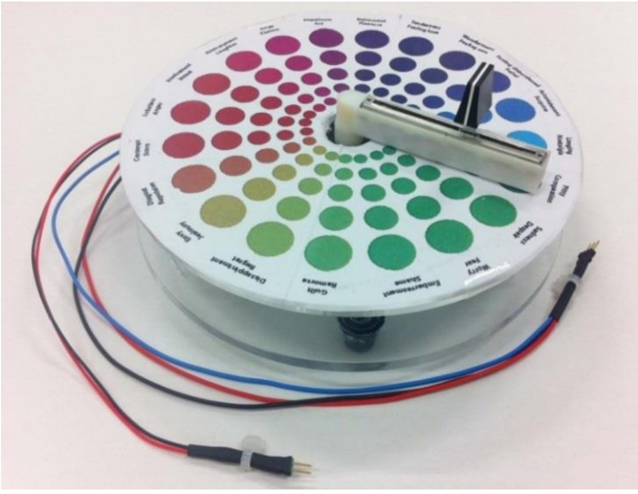

Figure 5: *The 2D labelling system (Haratian, 2016)*

4. CONCLUSION

A joystick which can sense biosignals and provide biofeedback to the user while interacting with computer is proposed in this paper. The collected biosignals are used for emotion recognition in real-time and a continuous manner to find out the quality and intensity of emotion using Machine Learning and Artificial Inteligence approaches. This intelligent embeded device has applications in different industries where the user is interacting with the machine through his/her hand while grabbing a joystick applied to aircraft, game and healthcare sectors.

5. REFERENCES

Journal article: Acampora, G., Cook, D. J., Rashidi, P. and Vasilakos, A. V., (2013) A Survey on Ambient Intelligence in Health Care. Proceedings of The IEEE. Institute of Electrical and Electronics Engineers, 101 (12), 2470-2494.

Journal article: Calvo, R. A. and D'Mello, S., (2010) Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. IEEE Transactions on Affective Computing, 1 (1).

Journal article: Greene, S., Thapliyal, H. and Caban-Holt, A., (2016) A Survey of Affective Computing for Stress Detection: Evaluating technologies in stress detection for better health. IEEE Consumer Electronics Magazine, 5 (4), 44.

Journal article: Gunes, H. and Pantic, M., (2010) Automatic, Dimensional and Continuous Emotion Recognition. International Journal of Synthetic Emotions, 1 (1), 68.

Journal article: Jenke, R., Peer, A. (2018) A cognitive architecture for modeling emotion dynamics: Intensity estimation from physiological signals" Cognitive Systems Research, Vol. 49, pp. 128-141, 2018.

Journal article: Kim, J. and André, E., (2008) Emotion recognition based on physiological changes in music listening. IEEE Transactions On Pattern Analysis And Machine Intelligence, 30 (12), 2067-2083.

Journal article: Klaus, R. S., (2005) What are emotions? And how can they be measured? Social Science Information, 44 (4), 695-729.

Journal article: Kulic, D. and Croft, E. A., (2007) Affective State Estimation for Human-Robot Interaction. IEEE Transactions on Robotics, 23 (5), 991-1000.

Journal article: Nardelli, M., Valenza, G., Greco, A., Lanata, A. and Scilingo, E. P., (2015) Recognizing Emotions Induced by Affective Sounds through Heart Rate Variability. IEEE Transactions on Affective Computing, 6 (4), 385.

Journal article: Nicolaou M., Gunes H., Pantic M., (2011) Continiuse Prediction of spontaneouse affect from multiple cues and modalities in valance-arousal space" IEEE Transactions on Affective Computing, 2(2)

Journal article: Picard, R. W., Vyzas, E. and Healey, J., (2001) Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. IEEE Transactions on Pattern Analysis & Machine Intelligence, 23 (10), 1175.

Journal article: Russell, S. J. and Norvig, P., (2003) Artificial intelligence : a modern approach [Non-fiction]. Upper Saddle River, N.J, Prentice Hall, 2nd ed.

Journal article: Santos, O. C., Uria-Rivas, R., Rodriguez-Sanchez, M. C. and Boticario, J. G., (2016) An Open Sensing and Acting Platform for Context-Aware Affective Support in Ambient Intelligent Educational Settings. IEEE Sensors Journal, 16 (10), 3865.

Journal article: Scherer, K. R., 2013. The GRID meets the Wheel: Assessing emotional feeling via self-report1 Oxford: Oxford University Press.

Journal article: Valenza, G., Lanata, A. and Scilingo, E. P., (2012) The Role of Nonlinear Dynamics in Affective Valence and Arousal Recognition. IEEE Transactions on Affective Computing, 3 (2), 237.

Journal article: Wanhui, W., Guangyuan, L., Nanpu, C., Jie, W., Shangguan, P. and Wenjin, H., (2014) Emotion Recognition Based on Multi-Variant Correlation of Physiological Signals. IEEE Transactions on Affective Computing, 5 (2), 126.

Book: Bishop, C. M., (2006). Pattern recognition and machine learning, New York : Springer, c2006.

Conference: Haratian, R. Timotijevic, T. (2018), On-body Sensing and Signal Analysis for User Experience Recognition in Human-Machine Interaction submitted to International Conference on Frontiers of Signal Processing

Technical document: Haratian, R. (2016), Dynamic Emotion Quality and Intensity Recognition through Physiological Signals with Application in Human Robot Interaction. Bristol Robotic Laboratory.

Internet source: Kreativebomb.com, (2011) Retro Joystick [online] available from http://www.kreativebomb.com/blog/2011/01/25/retro-joystick/ [Accessed 1 Dec. 2016].